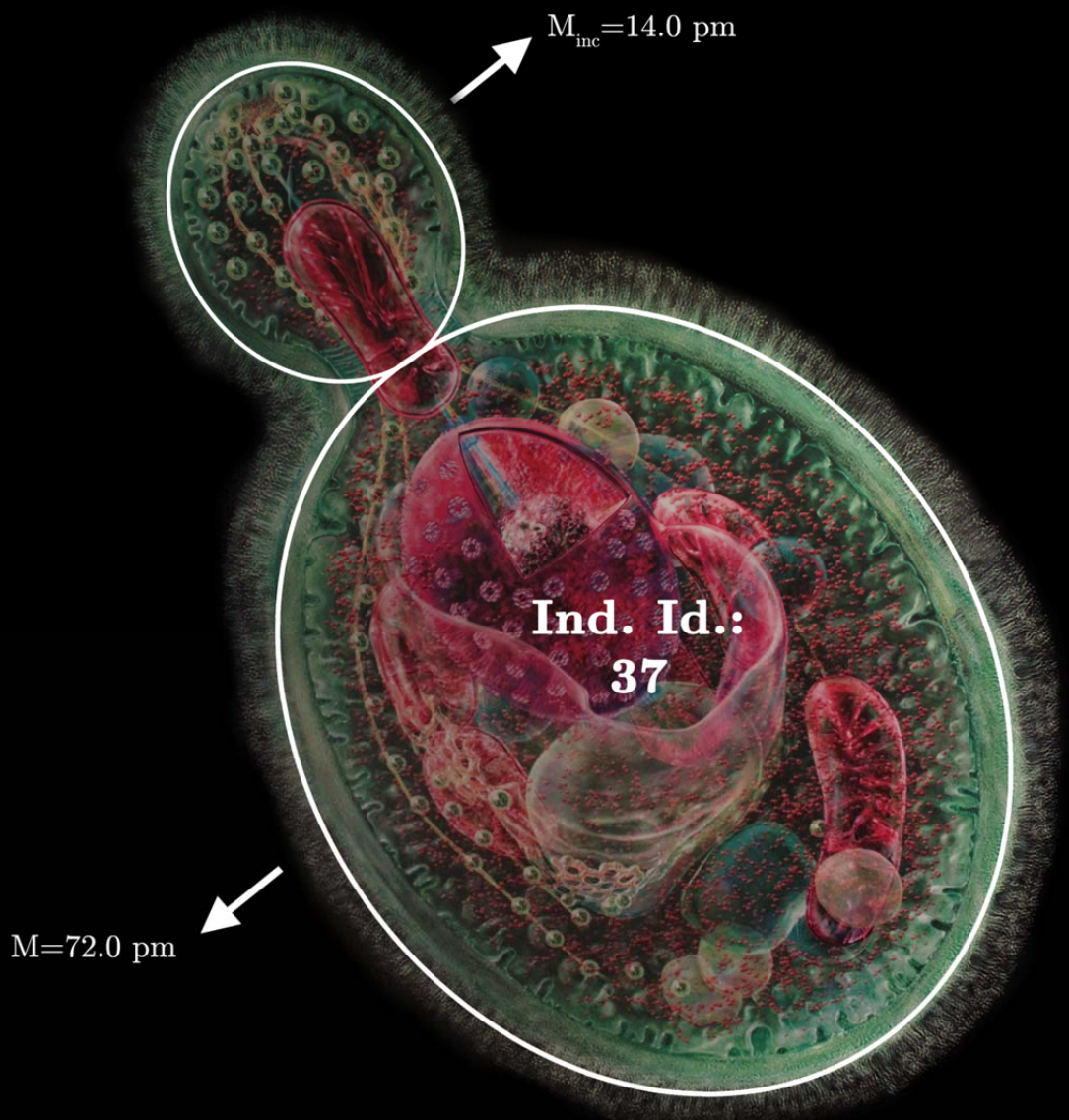
PhD THESIS

Individual-based observations and individual-based simulations to study *Saccharomyces cerevisiae* cultures



Xavier Portell Canal

Universitat Politècnica de Catalunya
Departament d'Enginyeria Agroalimentària i Biotecnologia
Programa de doctorat en
Tecnologia Agroalimentària i Biotecnologia

Individual-based observations and individual-based simulations to study *Saccharomyces cerevisiae* cultures

by
Xavier Portell Canal

Thesis submitted for the degree of
DOCTOR OF PHILOSOPHY
at the
Universitat Politècnica de Catalunya

Supervisors:

Anna Gras Moreu
Departament d'Enginyeria Agroalimentària i Biotecnologia

Marta Ginovart Gisbert
Departament de Matemàtica Aplicada III

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Departament d'Enginyeria Agroalimentària
i Biotecnologia

Castelldefels. September, 2014

Xavier Portell Canal
Tesi doctoral de la Universitat Politècnica de Catalunya
Setembre de 2014

Imatge de portada

La imatge utilitzada per a l'elaboració de la portada apareix originàriament a l'exposició anomenada "From Another Kingdom" (D'un Altre Regne) del National Botanic Garden of Wales (del 19 de març del 2011 al 28 de febrer del 2014). Disponible a flickr.com/col_and_tasha; National Botanic Garden of Wales sota llicència **CC-BY-2.0**.

Realització de la coberta

Núria Sallés Hipólito

Correcció ortogràfica

Sally Poulson (anglès)
Isaac Guilà i Puig (català i castellà)

"Science is what we understand well enough to explain to a computer. Art is everything else we do."

– Donald E. Knuth

Agraïments

Després de molt temps i d'haver-me d'aguantar la respiració en diverses ocasions, finalment ha arribat el moment. Permeteu-me que ho confessi: més d'un cop i dos he pensat que no em seria necessari escriure aquestes línies, però estic content de no haver-la encertada. Abans de començar, però, m'agradaria puntualitzar que, com diuen en castellà, *no están todos los que son, pero sí son todos los que están.*

És de justícia que comenci per les meves directores de tesi, la Dra. Marta Ginovart i la Dra. Anna Gras, de les quals he après moltes coses durant aquest temps i és obvi que aquesta tesi no hauria estat possible sense elles. Agraeixo, igualment, el suport cientificoeconòmic rebut per part del grup de recerca de Modelització i Simulació Discreta de Sistemes Biològics (MOSIMBIO) i el suport econòmic de la Universitat Politècnica de Catalunya. I molt especialment l'ajuda que m'han ofert dos dels "MOSIMBIOs", ara ja doctors, la Dra. Clara Prats i el Dr. Jordi "Tito" Ferrer: gràcies per deixar que m'aprofités de la vostra experiència com a doctorands i com a científics. *El siguiente serás tú, Pablo, gracias por la ayuda brindada y por tu interés en la cultura que, aunque temporalmente, tiene el placer de acogerte. Como decimos en catalán, y si me permites un consejo:* paciència i una canya!

Gràcies a tots els coautors de les contribucions que m'han permès portar aquest document a bon port. M'agradaria tenir unes paraules per als experimentals del grup, el catedràtic Josep "Temi" Vives-Rego i la Dra. Rosa Carbó: ha estat un plaer treballar amb tu, "Temi", la teva actitud i el teu *savoir-faire* són d'agrair i s'encomanen (o almenys, a risc de semblar egoista, això vull creure!). Infinita paciència és el que t'ha calgut a tu, Rosa, per aguantar totes les meves consultes sobre microbiologia. Si les circumstàncies ho permeten i no és molt abusar, espero poder continuar consultant-te el que em calgui durant els propers anys.

Es mereixen un bon agraïment, també, els estudiants de TFC/TFG amb els quals he pogut col·laborar i als quals la seva valuosa feina ha proporcionat dades per a la tesi: Marta Vias, Adrián Pérez i Julen Rivas. Molta sort en la vostra vida professional!

I would like to thank all the people from the Institute of Food Research (Norwich, UK) Foodborne Bacterial Pathogens research group for their help and assistance during my stay in Norwich. I am particularly indebted to: Dr Gary Barker, Dr Moisés Silbert, Mr Mark Fernandez, Dr Natalia Gómez, and Dr Pradeep Malakar. I would like to express my sincerest gratitude to the late Dr Moisés Silbert, a fellow "MOSIMBIO", whose warm welcome to Norwich made me feel completely at home. Rest in peace Moisés.

I tampoc no podien faltar aquí unes paraules per als "Marcians" del Campus. A l'equip de navegació (Chus Melo i Lluís Monfort), al "comandant científic" (Jordi Gutiérrez) i a tots els capitans (disculpeu que no n'escrigui tots els noms, però heu set molts i molt bons!). Crec, de debò, que hem fet una feina excel·lent!

Moltes gràcies als companys i als amics de l'ESAB (o "ESAB junkies", perquè aguantar "la que està caient" ha de ser cosa d'addicció) per les estones compartides: Clara Almansa, Sheila

Benítez, Joan Casals, Dani Fenero, Olga Gener, Ari Giné, Cristina Gonzalez, Maria Julià, Jordi Llop, Jordi Llorens, Graciela Marando, Íngrid Masaló, Miquel Masip, Marçal Plans, Míriam Pocurull, Ana Rivera, Aurora Rull, Joan Simó i molts d'altres amb els quals també he tingut el plaer de coincidir. Si hi ha un fet que lamento profundament és no haver pogut compartir més amb vosaltres.

A la secció de l'ESAB del departament de Matemàtica Aplicada III. Don Pelayo, Senyor Jarauta, ara sí que sí! Els agraïments més sincers i merescuts per a tu, Mònica Blanco, professora, la meva consellera (i psicòloga!), amiga i la millor companya de despatx que puc arribar a imaginar. Totes les paraules d'agraïment empal·lideixen davant la gesta encomanada, de manera que, simplement et dic MIL GRÀCIES!

Per acabar, un record per a la meva família: pares, germà Gerard, avis, tiets i cosins, cosines i cosinetes. A qui tinc més per agrair és a la meva mare, la Maria Teresa Canal, ja que de fa temps li ha tocat ser el puntal de la família en molts aspectes i és clar que sense ella no hauria arribat fins aquí. Voldria també recordar especialment els que ja no hi són: el meu pare, el Josep Portell Saus (sempre que algú em diu "Senyor Portell" em recordo de tu, pare!), els avis paterns (el Climent i la Remei) i els materns (el Josep i l'Encarnació), procuro recordar tot el que vaig aprendre de vosaltres; i, finalment, el tiet Francisco, un home temperat i assenyat com pocs que malauradament també ens va deixar.

No voldria finalitzar sense unes paraules d'agraïment per a la Meritxell Vancell: tu vas haver d'aguantar estoicament una gran part dels inconvenients derivats d'aquesta tesi. Ara que les circumstàncies de la vida han fet que ja no continuem junts vull que sàpigues que sempre tindràs un lloc al meu cor. Ho sento.

Resum

El *Saccharomyces cerevisiae* és un dels llevats que gaudeix de més significació econòmica, social i per a la salut humana. Depenent de les condicions experimentades, el llevat *S. cerevisiae* pot créixer mitjançant un metabolisme fermentatiu, respiratori o respirofermentatiu. La formació de cicatrius, una divisió desigual, una vida replicativa limitada i un increment de la mida de la cèl·lula amb l'edat replicativa són característiques individuals d'aquest llevat que afecten el comportament dels bioprocessos. Aquestes característiques incrementen la complexitat dels models predictius i dificulten, per tant, la seva inclusió en un model continu de manera realista. No obstant això, un model basat en l'individu sí que és capaç d'acomodar tota aquesta complexitat en un únic model computacional. Una vegada implementat, un model basat en l'individu ha de ser parametritzat, calibrat i la seva adequació ha de ser avaluada. Tots aquests processos requereixen idealment un gran nombre d'observacions experimentals, tant individuals com a nivell del sistema estudiat. L'objectiu general de la tesi present és avançar en el desenvolupament d'una metodologia basada en l'individu per estudiar sistemes microbians conduïts pel llevat *S. cerevisiae*.

Primerament s'avalua l'adequació de INDISIM-YEAST, un model basat en l'individu, ja existent, focalitzat en un llevat genèric. Es verifica i s'avalua la diversitat del *S. cerevisiae* en observacions experimentals orientades a l'individu en diferents condicions de creixement i en diversos estadis de la corba de creixement de la població. Això permet obtenir observacions basades en l'individu molt valuoses a l'hora de donar suport a la metodologia desitjada.

Es desenvolupa i s'implementa en Fortran 90 INDISIM-*Saccha*, un model quantitatiu basat en l'individu i focalitzat en el creixement fermentatiu (anaerobi) del *S. cerevisiae*. El model desenvolupat és parametritzat, calibrat, la seva adequació és avaluada i és utilitzat per estudiar *in silico* la producció d'etanol mitjançant experiments virtuals. El procés de calibratge, l'obtenció i l'anàlisi de les dades dels experiments virtuals s'han realitzat utilitzant el programari estadístic R. L'adequació del model s'avalua testejant diferents prediccions del model a nivell de sistema (corbes de disminució de la glucosa i de creixement de la població) i a nivell de la cèl·lula individual (evolucions temporals de la fracció de cèl·lules gemades, de la distribució d'edats genealògiques i de la distribució dels diàmetres cel·lulars). Les observacions del diàmetre de les cèl·lules individuals obtingudes a la tesi present juguen un paper significatiu en aquesta avaluació. Els resultats dels experiments virtuals suggereixen que les diferències en la distribució de mides cel·lulars poden afectar dràsticament l'evolució i la productivitat de les fermentacions i suggereixen una caracterització rutinària de l'inòcul a la indústria biotecnològica.

L'INDISIM-*Saccha* també és adaptat per tenir en compte el creixement aeròbic del *S. cerevisiae* i és contrastat mitjançant dos assajos experimentals amb dos nivells d'oxigen al medi. Els resultats preliminars de la simulació denoten que aquesta aproximació també té el potencial de reproduir cultius discontinus aerobis del *S. cerevisiae*. Això representa un pas endavant cap a l'obtenció d'un model basat en l'individu que tingui en compte tot el conjunt

d'alternatives metabòliques experimentades pel *S. cerevisiae*.

Finalment, aquesta tesi també dissenya i implementa INDISIM-YEAST-NL en l'ambient de programació lliure anomenat NetLogo per tal de comunicar de manera eficient, d'incrementar l'accessibilitat i d'afavorir l'ús de la metodologia INDISIM-*Saccha*. La implementació d'aquest model simplificat amb NetLogo posa les bases per a una comprensió més alta de la metodologia desenvolupada, i dels models microbians basats en l'individu en general, i facilitarà futures interaccions amb usuaris potencials de l'INDISIM-*Saccha*.

Paraules clau: llevat, model basat en l'individu, parametrització, calibratge, adequació del model, experiment virtual, anàlisi elèctric de partícules, difracció de la llum, NetLogo.

Resumen

Saccharomyces cerevisiae es una de las levaduras con mayor significación económica, social y para la salud humana. Dependiendo de las condiciones experimentadas, *S. cerevisiae* puede crecer mediante un metabolismo fermentativo, respiratorio o respirofermentativo. La formación de cicatrices, una división desigual, una vida replicativa limitada y el incremento del tamaño de la célula con la edad replicativa son características individuales de esta levadura que afectan el comportamiento de los bioprocesos. Estas características incrementan la complejidad de los modelos predictivos, dificultando, por tanto, su inclusión en un modelo continuo de manera realista. No obstante, un modelo basado en el individuo sí es capaz de acomodar toda esta complejidad en un único modelo computacional. Una vez implementado, un modelo basado en el individuo debe ser parametrizado, calibrado y su adecuación debe ser evaluada. Todos estos procesos requieren idealmente un gran número de observaciones experimentales, tanto a nivel individual como a nivel del sistema estudiado. El objetivo general de la presente tesis es avanzar en el desarrollo de una metodología basada en el individuo apta para estudiar sistemas microbianos conducidos por la levadura *S. cerevisiae*.

Primeramente se evalúa la adecuación de INDISIM-YEAST, un modelo basado en el individuo, ya existente, centrado en una levadura genérica. Se verifica y evalúa la diversidad en observaciones experimentales orientadas al individuo de *S. cerevisiae* en diferentes condiciones de crecimiento y en diferentes estadios de la curva de crecimiento de la población. Esto permite obtener observaciones basadas en el individuo muy valiosas a la hora de apoyar la metodología deseada.

Se desarrolla e implementa en Fortran 90 INDISIM-*Saccha*, un modelo cuantitativo basado en el individuo focalizado en el crecimiento fermentativo (anaerobio) de la levadura *S. cerevisiae*. El modelo es parametrizado, calibrado, su adecuación es evaluada y es utilizado para estudiar *in silico* la producción de etanol mediante experimentos virtuales. El proceso de calibrado, la obtención y el análisis de los datos experimentales virtuales se realiza utilizando el software estadístico R. La adecuación del modelo se evalúa testeando diferentes predicciones del modelo a nivel de sistema (curvas de disminución de la glucosa y de crecimiento de la población) y a nivel de célula individual (evoluciones temporales de la fracción de células gemadas, de la distribución de edades genealógicas y la distribución del diámetro celular). Las observaciones del diámetro de las células individuales obtenidas en la presente tesis desempeñan un papel significativo en esta evaluación. Los resultados de los experimentos virtuales sugieren que diferencias en la distribución de tamaños celulares pueden afectar drásticamente la evolución y productividad de las fermentaciones y sugieren una caracterización rutinaria del inoculo en la industria biotecnológica.

INDISIM-*Saccha* también es adaptado para tener en cuenta el crecimiento aerobio de *S. cerevisiae* y contrastado mediante dos ensayos experimentales con dos niveles de oxígeno en el medio. Los resultados preliminares de simulación sugieren que esta aproximación también

tiene el potencial de reproducir cultivos discontinuos aerobios de *S. cerevisiae*. Esto representa un paso más hacia la obtención de un modelo basado en el individuo que tenga en cuenta todo el conjunto de alternativas metabólicas experimentadas por *S. cerevisiae*.

Finalmente, la presente tesis también diseña e implementa INDISIM-YEAST-NL en el ambiente de programación libre NetLogo para comunicar de manera eficiente, incrementar la accesibilidad y favorecer el uso de la metodología INDISIM-*Saccha*. La implementación de este modelo simplificado en NetLogo sienta las bases para una mayor comprensión de la metodología desarrollada, y los modelos microbianos basados en el individuo en general, a la vez que facilitará futuras interacciones con usuarios potenciales de INDISIM-*Saccha*.

Palabras clave: levadura, modelo basado en el individuo, parametrización, calibración, adecuación del modelo, experimento virtual, análisis eléctrico de partículas, difracción de la luz, NetLogo.

Abstract

Saccharomyces cerevisiae is one of the yeasts with major economic, social, and health significance in human culture. Depending on the growth conditions experienced by the cell, *S. cerevisiae* growth can proceed via fermentative, respirative, or respirofermentative metabolism. Scar formation, unequal division, a limited replicative lifespan, and increase in cell size commensurate with the cell's replicative age are individual characteristics of this yeast affecting the performance of bioprocesses. These characteristics increase the complexity of predictive models and introducing them with ease into a continuous model is not realistic. Nevertheless, an individual-based model is able to accommodate this complexity in a single computational model. Once an individual model is implemented, it has to be parameterized, calibrated, and its adequacy assessed. All these processes ideally require a high number of both individual and system-level experimental observations. The aim of the present thesis is to advance the development of an individual-based methodology to tackle the study of microbial systems driven by the relevant yeast *S. cerevisiae*.

The adequacy of INDISIM-YEAST, an existing individual-based model of a generic budding yeast, is first assessed. In order to obtain valuable individual-based observations to support the desired individual-based methodology, the diversity of *S. cerevisiae* in experimental individually-oriented observations under different growth conditions and at different stages of the growth curve is verified and assessed.

A quantitative individual-based model focusing on the fermentative (anaerobic) growth of the yeast *S. cerevisiae* has been designed, implemented in Fortran 90, and termed INDISIM-*Saccha*. The developed model is parameterized, calibrated, its adequacy evaluated, and used to assess *in silico* ethanol production by means of virtual experiments. The calibration procedure, and the performance and analysis of the data from the virtual experiments is undertaken using the statistical programming language R. The model adequacy is assessed by testing several model predictions both at a system level (glucose depletion, population growth curves) and single-cell level (fraction of budded cells, genealogical age distribution, and cell diameter distribution evolutions). Individual cell diameter observations obtained within the present thesis play a significant role in this assessment. Results of the virtual experiments suggest that differences in cell size distribution can drastically affect the performance and productivity of fermentations, and encourage routine characterization of the inocula in the biotechnological industry.

INDISIM-*Saccha* is also adapted to take into account the aerobic growth of *S. cerevisiae* and contrasted with two experimental trials with different oxygen levels in the medium. The preliminary simulated results achieved with the model suggest that the approach also has the potential for reproducing aerobic batch cultures of *S. cerevisiae*. This represents a further step in obtaining a microbial individual-based model to account for the whole set of metabolic alternatives experienced by *S. cerevisiae*.

In order to communicate efficiently, increase accessibility, and favour usability of the INDISIM-*Saccha* methodology developed, the present thesis also designs and implements INDISIM-YEAST-NL in the freely available programming environment NetLogo. The implementation of this streamlined model in NetLogo lays the foundations for a deeper understanding of the developed methodology and microbial individual-based models in general, and will facilitate future interactions with potential users of INDISIM-*Saccha*.

Keywords: yeast,individual-based model, parameterization, calibration, model adequacy, virtual experiment, flow cytometry, electrical particle analysis, light diffraction, NetLogo.

Contents

List of Acronyms

AbM Agent-based Model

DAPI diamidino-2-phenylindole

FISH Fluorescent *In Situ* Hybridization

FITC Fluorescein IsoThioCyanate

IbM Individual-based Model

IbO Individual-based Observation

INDISIM INDividual DIScrete SIMulation

INDISIM-COMP INDividual DIScrete SIMulation of the COMPosting process

INDISIM-*Paracoccus* INDividual DIScrete SIMulation of *Paracoccus denitrificans*

INDISIM-RBC INDividual DIScrete SIMulation of Red Blood Cells

INDISIM-*Saccha* INDividual DIScrete SIMulation of *Saccharomyces cerevisiae*

INDISIM-SOM INDividual DIScrete SIMulation of Soil Organic Matter

INDISIM-YEAST INDividual DIScrete SIMulation of a generic budding YEAST

INDISIM-YEAST-NL Streamlined version of INDISIM-*Saccha* implemented in NetLogo

ODD Overview, Design concepts, and Details protocol

μIbE microbial Individual-based Ecology

μIbM microbial Individual-based Model

Chapter 1

Introduction

1.1 The yeasts

Yeasts are of major economic, social and health significance in human culture, often described has being the mankind's oldest "domesticated" organism with the brewing of beer being probably the world's first biotechnology. An overview of the branches of yeast biotechnology can be seen in Fig. 1.1.

Although less ubiquitous than bacteria, yeasts are widespread in nature, being found in terrestrial, aquatic, and aerial environments. They are non-motile, relying on aerosols, animal vectors and human activity for dispersal; and strictly chemoorganotrophic microorganisms, thus requiring fixed, organic carbon sources for growth. The presence of a yeast species in a particular niche is determined by the carbon sources metabolizable by it (Walker, 1998). Although most yeasts are mainly associated with plants or animals, some species are able to cause infections in humans and other animals, but most act opportunistically in immune compromised individuals.

Yeasts are a group of eukaryotic fungi with a well-defined cell wall comprised primarily of the D-glucose polysaccharide β-(1–3)-glucan with growth being either entirely unicellular or a combination of hyphal and unicellular reproduction under particular growth conditions (Walker, 1998). Yeasts belong to two separate phyla of the kingdom Fungi: the Ascomycota phylum, and the Basidyomicota phylum. They have divers evolutionary origins and biochemical properties and consequently can be found within several subphyla or classes of the Ascomycota and Basidyomicota phyla. Sexual reproduction of yeasts involves the formation of internally formed ascospores or externally formed basidiospores. Although the concept of yeast has not always followed logical biological lines (Lachance, 2001), they have been defined as "ascomycetous or basidiomycetous fungi that reproduce vegetatively by budding or fission, and that form sexual states which are not enclosed in fruiting bodies" (Boekhout and Kurtzman, 1996).

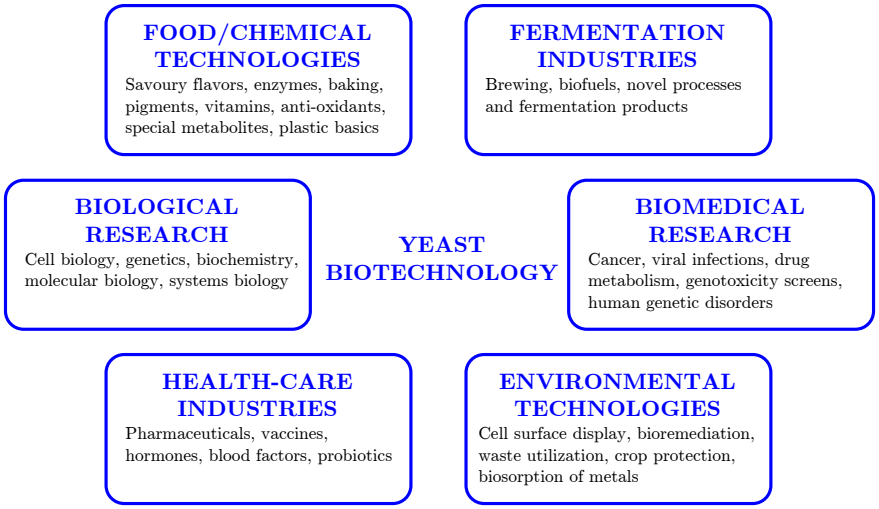


Figure 1.1: Branches of yeast biotechnology. Re-elaborated from Feldmann (2012).

1.2 The relevant yeast *Saccharomyces cerevisiae*

The yeast *Saccharomyces cerevisiae* (Table 1.1) is, without doubt, one of the most notable and studied microorganisms. Other common names for this species are brewer's yeast (also top fermenting or ale yeast), baker's yeast, budding yeast, and is even used as a synonymous for yeast in everyday language (Feldmann, 2012; Kurtzman and Fell, 1998). The scientific name "Saccharomyces" derives from the Greek and means "sugar fungus" ("Saccharo" = sugar and "myces" = fungus), while "cerevisiae" comes from the Latin and means "of beer", and stems from Ceres, the Roman God of the crops. Although present for thousands of years, it is only relatively recently that it has been identified and put into use. Yeasts (in all probability, *S. cerevisiae*) were used for brewing beer in Sumeria and Babylonia around 6000-7000 BC; and, at the same time, *S. cerevisiae* strains were used to produce wine in Georgia (Feldmann, 2012). The use of yeasts for dough leavening by the production of carbon dioxide in Egypt occurs much later, ca. 4000 BC (Tuite and Oliver, 1991).

Table 1.1: Taxonomic classification of the yeast *S. cerevisiae*.

Taxonomic category	
Subdivision	*Ascomycotina*
Family	*Saccharomycetaceae*
Subfamily	*Saccharomycetoideae*
Genera	*Saccharomyces*
Species	*Cerevisiae*

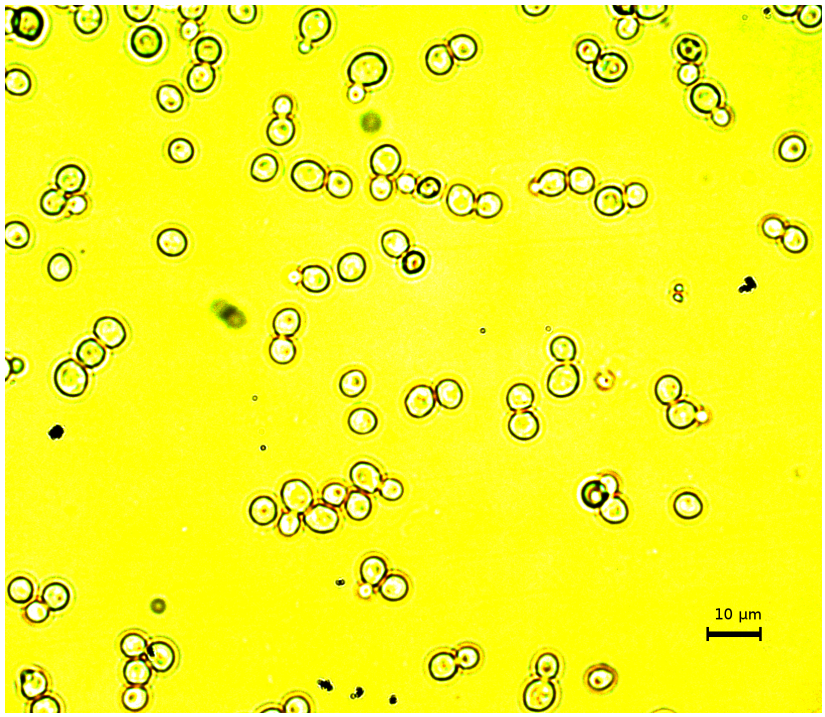


Figure 1.2: Unstained aerobically grown stationary phase (27 h) *S. cerevisiae* var. *bayanus* cells under the optical microscope (x500).

1.2.1 Overview of the biology of *Saccharomyces cerevisiae*

S. cerevisiae is a unicellular, non-motile microorganism that propagates easily and relatively quickly, is highly adaptable to large-scale production, and possesses the cellular organization and advantages of cell compartmentalization of multicellular eukaryotes. When observed under a microscope, *S. cerevisiae* cells are seen to be ovoid or ellipsoidal structures surrounded by a rather thick cell wall (Fig. 1.2). Mean values for the large diameter range between 5 and 10 μm, and for the small diameter between 1 and 7 μm (Feldmann, 2012). Mean cell size of *S. cerevisiae* also increases as the cells age. Brewing strains of *S. cerevisiae* are usually bigger than laboratory (or "Scientific") strains.

Laboratory strains of *S. cerevisiae* can exist either as haploid or as diploid cells. Both forms may exist as stable cultures and undergo repeated rounds of vegetative growth and mitosis. Haploid cells of opposite mating types can fuse, allowing formation of a diploid cell. Diploid cells undergo meiosis and spore formation when starved of nitrogen and in the presence of a poor carbon source (e.g., acetate), which results in four haploid spores contained in the wall of the mother cell (the ascus). Spores in rich nutrient conditions will germinate and start a new haploid cycle.

Unlike bacteria and other yeasts, the vegetative proliferation of *S. cerevisiae* is by asymmetrical budding, that is, the detached bud (daughter cell) is smaller than the parent (mother)

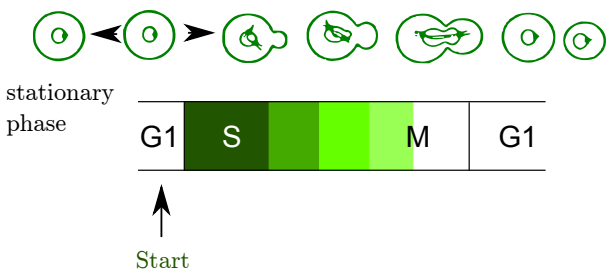


Figure 1.3: Phases of the *S. cerevisiae* cell cycle. This image shows the position of the haploid cell relative to phases of the cell cycle. Cells bud once "Start" is passed, with the emergence of the bud representing a simple morphological marker of the entry into the S phase. No simple morphological or biochemical marker of the entry into mitosis exists. Redrawn from Walker (1998).

cell. The cell cycle of this yeast consists of four phases regulated in a way similar to that of the larger eukaryotes, the G1, S, G2, and M phases (Fig. 1.3). During G1 the cell grows until the underlying molecular mechanisms (see Dickinson and Schweizer, 2004) trigger the "Start" commitment point to the cell cycle seen by bud emergence. At "Start" a yeast cell integrates environmental and internal signals (nutrient availability, presence of pheromone, attainment of critical size, metabolic machinery status), and decides to enter into a new cell cycle (Alberghina *et al.*, 2012). The initiation of DNA synthesis is triggered at "Start". In *S. cerevisiae*, the mitotic spindle forms very early and the S and M Phases overlap (Fig. 1.3), resulting in an indistinct G2 phase in normal cells (Walker, 1998). Once in M phase, the newly synthetized chromosomes migrate to the bud, which is detached once it has attained a minimum size. The bud abscission leaves a chitin bud scar on the mother cell and a birth scar on the daughter cell. The number of bud scars can be used to follow the number of divisions an *S. cerevisiae* cell has undergone, i.e., its genealogical age (Fig. 1.4). *S. cerevisiae* cells are not able to produce new daughters indefinitely but each strain has its characteristic replicative lifespan, which is usually measured as the mean or median number of generations produced before it is no longer able to continue proliferating (Dickinson and Schweizer, 2004). This replicative limit is also known as the Hayflick limit. Besides the replicative lifespan, a chronological lifespan also exists which has been defined as a consequence of cumulative and irreversible damage to intracellular components during the extended stationary phase, compromising cell integrity, and leading to death and autolysis. Both lifespan forms have been believed to represent models of ageing in higher eukaryotes (Maskell *et al.*, 2003).

S. cerevisiae is able to grow on a wide range of carbon compounds but glucose is the preferred substrate because the presence of this hexose inhibits the utilization of other carbon sources (Dickinson and Schweizer, 2004). *S. cerevisiae* catabolizes the glucose via the Embden-Meyerhof pathway (glycolysis) leading to the production of two molecules of pyruvate which can be then fermented to ethanol and carbon dioxide (Eq. 1.1), or oxidized via the citric acid cycle (Krebs cycle) and electron transport chain to carbon dioxide and water with the formation of more ATP (Eq. 1.2). The occurrence of either one or other alternative is mainly

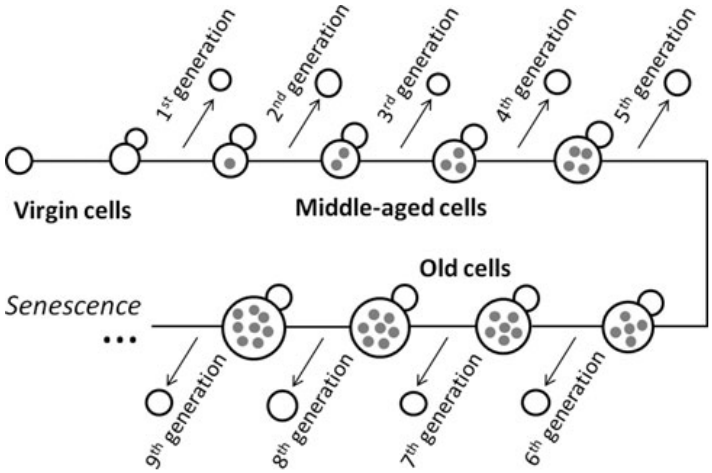


Figure 1.4: Progression from a newly emerged daughter cell (virgin cell) to senescence, moving through the different genealogical ages, which allows classification of the parent cells into middle-aged cells (or young parent cells) and old cells (or old parent cells). A daughter cell is formed at each step of cell division.

regulated by the glucose and oxygen availabilities.

$$C_6H_{12}O_6 \rightarrow 2\ C_2H_6O + 2\ CO_2 + 2\ ATP \tag{1.1}$$

$$C_6H_{12}O_6 + 6\ O_2 \rightarrow 6\ H_2O + 6\ CO_2 + 36\ ATP \tag{1.2}$$

S. cerevisiae is described as a Crabtree-positive yeast and, thus, under aerobic growth conditions with abundant glucose as energy source, fermentation predominates over respiration. A long-term and a short-term Crabtree effect can be found. A long-term Crabtree effect is the occurrence of aerobic fermentation under fully adapted, steady-state conditions at high growth rates. On the other hand, a short-term Crabtree effect is a sudden fermentative response under fully aerobic conditions upon addition of excess sugar to yeast that was not fermenting before the addition of the sugar. The long-term Crabtree effect has been explained in terms of a limited respiratory capacity of the yeast. The result of this, depending on the growth conditions experienced by the yeast, is that cell growth can continue via fermentative, respirative, or respirofermentative (a.k.a., oxido-reductive) metabolism. The pure fermentative or respirofermentative growth is exploited in the making of bread, beer, wine, and other alcoholic beverages while, for the production of yeast biomass, the respirative metabolism is desirable.

S. cerevisiae is able to grow anoxically in a suitable medium, which has been interpreted as indicating that respiration is absolutely optional for it (Alexander and Jeffries, 1990). In a non-suitable medium , however, *S. cerevisiae* requires anaerobic growth factors. This is due to the oxygen requirement of the *de novo* biosynthesis of sterols, unsaturated fatty acids, and other molecules (reviewed by Ishtar Snoek and Yde Steensma, 2007; Rosenfeld *et al.*, 2002).

1.2.2 The importance of *Saccharomyces cerevisiae* to man

For many thousands of years, man has benefited from using *S. cerevisiae*. Nowadays, this yeast is involved in myriad economic, social, and health-related products. Natural products include fermentation products and cell components, or the whole cell itself. Recombinant DNA technology, along with its ample history of use and accumulated knowledge of the microorganism, has made *S. cerevisiae* one of the preferred cell factories specifically engineered to produce and commercialize a great number of recombinant metabolites. Furthermore, this organism has been playing an important role in both applied and basic biological research.

Industrial *S. cerevisiae* fermentations contribute significantly to the economies of many countries. Strains of *S. cerevisiae* are involved in the manufacturing of many alcoholic beverages, including beer, wine, cider, sake, and distilled spirits and distilled spirits (whisky, vodka, cognac, gin, liqueurs, etc.). Beer and wine are leading fermentation products worldwide in terms of tons per year (Kurtzman and Fell, 1998).

Fermentation biofuels such as bioethanol and biobutanol are also used world-wide as a source of renewable energy. The vast majority of the bioethanol produced in the world uses *S. cerevisiae*. The genetic engineering of yeasts, mainly using *S. cerevisiae* as a recombinant host, is believed to be of particular interest in increasing supplies of bioethanol from plant biomass and the fermentation of lignocellulosic hydrolysates (Feldmann, 2012; Walker, 1998). Genetically manipulated *S. cerevisiae* has also been proposed for the forthcoming production of biobutanol as this seems likely to be a better replacement for gasoline than ethanol. *S. cerevisiae* also produces higher alcohols (a.k.a. fusel oils) and polyhydric alcohols, such as glycerol. Some higher alcohols can be removed during the production of potable spirits and sold as solvents (e.g. isoamyl alcohol in perfumes). Relevant polyhydric alcohols are glycerol and xylitol. Glycerol has been industrially produced in the manufacture of nitroglycerine-based explosives but is also used in the production of synthetic resins, pharmaceuticals, cosmetics and toothpastes.

S. cerevisiae biomass itself has a number of relevant industrial uses and it represents the largest bulk production of any single-celled microorganism in the world. When the yeast themselves have commercial value, they are called commercial yeasts, for instance, baker's yeast, brewer's yeast, and distiller's yeast; the former being the main biomass-derived product. Whole *S. cerevisiae* cells additionally have a range of new applications (Walker, 1998), including use as livestock growth factor (feed yeast), biotherapeutic agent, chemical "reagent", biocontrol agent, biosorbent/bioremediation agent, and biosensor.

Yeasts are nutritionally interesting; they are a rich source of proteins, nucleic acids, vitamins, and minerals, with negligible levels of triglycerides. In Europe and the United States each resident consumes about 2 kg of yeasts per year (Feldmann, 2012). The lysis of the whole cell releases peptides, amino acids, vitamins, and other cellular components. The insoluble components (cell wall or "hulls") are usually removed to make yeast extracts, which are mainly produced with brewer's and baker's yeast. Products from yeast lysis have characteristic savoury flavours and aromas and thus may be used as food additives. Some product examples

using yeast extracts are dried soups, sauces, meat and fish preparations, gravy granules, and flavoured potato snacks. Removed cell walls from the yeast extract manufacture has further potential uses (see Feldmann, 2012; Walker, 1998), including many pharmaceutical, and food additive (thickening agent, fat substitute) applications, as well as having probiotic activity.

Invertase, the enzyme that catalyses the hydrolysis of sucrose to glucose and fructose, comes from autolysis of baker's yeast and has many applications in the confectionery industry and in the production of invert sugar. Industrially useful chemicals produced from *S. cerevisiae* are the amino acids lysine, glutamic acid and methionine, and the sterol ergosterol.

A range of factors and characteristics (see Feldmann, 2012; Porro *et al.*, 2011; Walker, 1998) have made *S. cerevisiae* into a microbiological workhorse in the commercialization of industrial recombinant metabolites and especially for recombinant human proteins. Some of the main engineered products, secondary metabolites and end-fermentation products, produced by this microorganism include Isoprene derivatives (e.g., beta-carotene, Astaxanthin, Artemisin, Farnesene), pigments (flavonone), other valuable biocompounds (vit C, hydrocortisone, resveratrol polyketides), and organic acids (succinic acid, lactic acid). Recombinant protein production is a multi-billion dollar market allowing production of biopharmaceuticals and industrial enzymes. In 2009, 50 % of the biopharmaceutical proteins, which account for the majority of recombinant biotherapeutics, were produced by *E. coli* (30%) or *S. cerevisiae* (20%) host platforms (see Porro *et al.*, 2011). The approved protein products in yeast are obtained exclusively in *S. cerevisiae*, (reviewed by Ferrer-Miralles *et al.*, 2009). At the beginning of 2009, the list included hormones (insulin, insulin analogues, non-glycosylated human growth hormone somatotropin, glucagon), vaccines (hepatitis B virus surface antigen), and virus-like particles of the major capsid protein L1 of human papillomavirus type 6, 11, 16, 18, urate oxidase from *Aspergillus flavus*, granulocyte-macrophage colony stimulating factor, algumin, hirudin of *Hirudo medicanalis* and human platelets derived growth factor. Therapeutic areas addressed by most of the recombinant pharmaceuticals from yeast are infectious diseases and endocrine, nutritional and metabolic disorders. The greatest value provided by *S. cerevisiae* recombinant technology is the production of human insulin, which covers half the insulin needed by diabetic patients around the world (Feldmann, 2012).

S. cerevisiae is closely involved in both applied and basic research. It was adopted as a model system for laboratory study in the 1930s and has, since then, become a proven model eukaryote for molecular and cellular biology studies (Forsburg, 2005; Mager and Winderickx, 2005). These have allowed discoveries related not only to the cell cycle control and damage responses, but also to virtually every aspect of cell behaviour from chromosome segregation to protein secretion. Major advances in yeast technology have stimulated a new recovery of this model organism in screening for new biologically active compounds and the elucidation of drug-induced molecular mechanisms with medical and medicinal purposes. Additionally, yeasts can be used for investigating (human) disease-related proteins that have no apparent homologous counterpart in this organism, the so-called humanized yeast systems, holding great promise for the dissection of disease-related molecular processes and the discovery of new

medical compounds. Studies addressing neurodegenerative disorders mainly caused by protein misfolding, such as Huntington's Alzheimer's, and Parkinson's disease are clear examples of this use.

1.3 Individual-based experimental observations

The study of the microbial world is inherently affected by our limited observation capacity of the system itself, requiring observations at a microscopic scale (or lower) and with widely variable time scales according to the desired objective. This characteristic necessitates the close linking and subediting of the study of any microbial system with the technological advances made at any particular time. Until recently, the vast majority of microbial studies have been driven by means of bulk approaches where the mean behaviour or state of the population has been targeted. Such population or bulk approaches have allowed us to achieve an ample pool of interesting knowledge, but their accuracy is called into question by the realization that genetically identical bacterial cells in a well-mixed environment may show individually differing phenotypes (e.g., Blake *et al.*, 2003; Graumann, 2006; Wessel *et al.*, 2013). Individual microorganisms, even those in "clonal" populations, may differ widely from each other in terms of their genetic composition, physiology, biochemistry, and behaviour. This variability or heterogeneity has important consequences (see Brehm-Stecher and Johnson, 2004, and references therein), including antibiotic and biocide resistance, the productivity and stability of industrial fermentations, the efficacy of food preservatives, and the potential of pathogens to cause disease. Accordingly, it seems obvious that a complete understanding of the microbial world will undoubtedly require, at least, experimental observations at a lower level of observation than the population, i.e., Individual-based Observations (IbOs) from the individual or single-cell level.

Finally, it now seems that the stated bulk approach can at last be surpassed with the increase in new technologies for single-cell microbiology. Schematically, the single-cell experimentation hinges mainly on the development of single-cell isolation techniques (dilution, micromanipulation, flow cytometry and cell sorting, microfluidics, and compartmentalization techniques), and of a variety of culture-independent analysis techniques (e.g., fluorescence, cytometry, scanning probe microscopies and microspectroscopic methods). An overview of the current single-cell isolation techniques may be found in the reviews of Ishii *et al.* (2010) and Wessel *et al.* (2013), while the tools and technologies available for single-cell microbiology are reviewed by Brehm-Stecher and Johnson (2004) (see also the reviews of Ishii *et al.*, 2010; and Musat *et al.*, 2012). Even the DNA of an isolated cell may be amplified and sequenced, a field of microbiology that has been called "single-cell genomics" (Walker and Parkhill, 2008). All these existing technologies, cited in Kreft *et al.* (2013), are finally enabling microbiologists to do what plant and animal ecologists have been doing routinely: "observe who does what, when, where, and next to whom."

1.4 Microbial size distributions

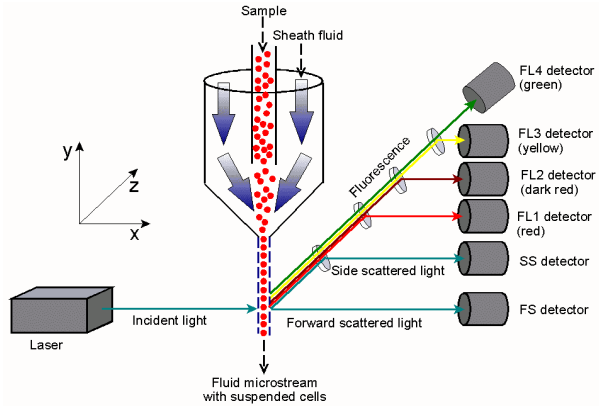
Size is the most fundamental aspect of cellular form, and the coupling of growth and division is the basis of cell size control. In yeast, this control occurs at two points: the boundaries between the G1 and S phases (a.k.a. "Start"), and between the G2 and M phases of the eukaryotic cell cycle. Although the cell-cycle regulatory network has been extensively described, the molecular mechanisms coupling cell growth to division, which finally controls cell size, are not completely understood (recently reviewed by Turner *et al.*, 2012). Although cell growth is more fundamentally related to the mass than the volume, the lack of methods to measure the individual cell mass has resulted in most of the research into cell size and growth being focused on volume (Godin *et al.*, 2010).

Cell size is an easily observable indication of the physiological state of a microbial population (Shuler *et al.*, 1972). Cell size measurements have been shown to be relevant in the study of the response of yeast cells subjected to different stresses: osmotic stress, heat stress, oxidative stress, hyperbaric stress and ethanol stress (see references in Tibayrenc *et al.*, 2010); and as an indicator to assess yeast vitality (see Walker, 1998) and viability (Tibayrenc *et al.*, 2010). Additionally, it is recognized that, although not sufficiently implanted, the control of even basic variables such as cell count, cell size distribution, and cell morphology may greatly improve the quality and stability of the biotechnological processes (Höpfner *et al.*, 2010). This, however, is encouraging the appearance of *in-situ* sensors (directly installed into the bioreactor) and affordable (offline) devices for bioprocess monitoring and control. Two recent reviews of the optical procedures, mainly focusing on *in-situ* microscopy systems, for the inline determination of cell count, cell size distribution, and other variables are provided by Bluma *et al.* (2010) and Höpfner *et al.* (2010). Examples of rapid, easy-to-use and affordable devices fitting the needs of the food industry include flow cytometers (e.g., EMD Millipore, Corp., 2013) and particle sizing and counting analysers (KGaA Merck, 2013).

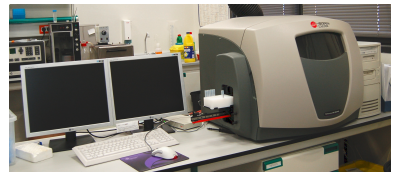
A variety of sizing methodologies exists. Forthcoming recently developed methods using interferometry (Popescu *et al.*, 2008; Wang *et al.*, 2011), which allows the dry mass of single cells to be measured with femtogram sensitivity, and microchannel resonators, which allows the buoyant mass of single cells to be measured (Bryan *et al.*, 2010; Godin *et al.*, 2010), will probably play an important role in the near future. Nevertheless, nowadays, in the biotechnological context the methods most frequently used to obtain cell size distributions are: flow cytometry, image cytometry, electric particle analysis, and light diffraction. These methodologies are further discussed below.

1.4.1 Flow cytometry

Flow cytometry is a versatile promising technique due to its high-throughput capacity while maintaining a single-cell level capability. Cells in a liquid sample are passed individually in front of an intense light source and data from light-scattering and/or fluorescence measurements are collected and saved (Fig. 1.5). In light-scattering measurements, the incident light scattered



(a)



(b)

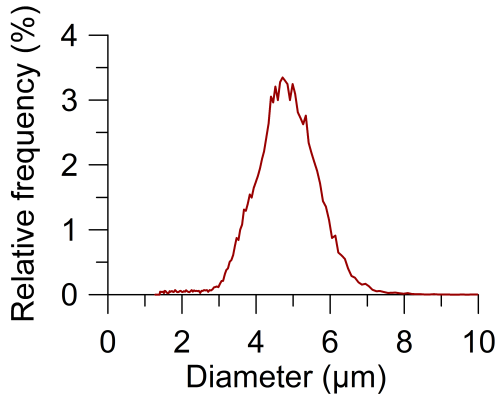
Figure 1.5: Flow cytometer operation outline (a) and the Cytomics 500 MPL flow cytometer of Coulter®Corporation (b). In (a): red circles represent microorganisms suspended in the fluid stream; blue arrows symbolize the incident light beam and the scattered light collected at different angles by the forward scatter (FS) and side scatter (SS) detectors; and red, dark red, yellow, and green arrows represent four different fluorescence emissions that are collected by corresponding detectors.

from one cell is collected at two different angles: a narrow forward angle (forward scatter) and a proximal right angle from the light beam (side scatter). Forward scatter is a complex parameter as it varies not only with cell size, but also with cell shape, refractive index, and the number of intracellular dielectric interfaces. Side scatter is also a complex parameter. This signal is thought to indicate variations in cell surface structure or internal structure, usually referred to in the literature as "cellular granularity" (Allman *et al.*, 1990). The combination of flow cytometry with fluorescent dyes contributes to the high potential of the technique, allowing determination (Alberghina and Porro, 1993) of a number of both "functional" (e.g., redox state, membrane integrity, and intracellular pH) and "structural" parameters (e.g., DNA content, DNA base ratio, total protein, basic proteins, and antigens). A relevant fluorescent dye is Fluorescein IsoThioCyanate, or FITC, which allows the cell's protein content to be determined, i.e. total fluorescence after FITC staining, a currently used measure of the size of the cells. Flow cytometric size distributions (cell volume or cell protein content) are drawn as a frequency function that reports the signal intensity of the individual cells on the abscissa, and the relative number of cells on the ordinate. As this is an indirect method of obtaining the size of the cells, in order to determine the biovolume (e.g., μm^3) from the described distributions (relative intensity), a calibration step using an appropriate standard has to be performed (e.g., Foladori *et al.*, 2008).

1.4.2 Electrical particle analysis

Electrical particle analysers are based on the principle of resistive pulse sensing (also referred to as electrical resistance method). Resistive pulse detection is based on the transient resist-

(a)



(b)

Figure 1.6: The Coulter®Counter electric particle size analyser Multisizer II (a), and a size distribution of aerobically grown stationary phase cells of a *S. cerevisiae* wine yeast strain recorded with the stated device (b).

ance modulation when a particle/cell passes through a small sensing aperture and displaces a volume of electrolytes equivalent to its own volume. The passage of a non-conducting particle suspended in a conducting solution through the small aperture results in a change in resistance across the aperture, measured as a voltage change (Shuler *et al.*, 1972). Since cells are constrained in a path or flowing stream and pass one at a time through a sensor that analyses the property of interest (voltage change), electric particle analysis (also referred as electronic particle analysis) can be considered as a flow cytometric analysis technique (Kell *et al.*, 1991; Turner *et al.*, 2012). Electrical particle analysis is a fast, direct (it provides a cell biovolume distribution, see Fig. 1.6), and accurate technique. However, unlike optically-based flow cytometry, this methodology is limited to cell counting and obtaining cell-volume distribution.

1.4.3 Light diffraction

Light diffraction (laser diffraction or laser light scattering) is a fast, accurate and first principle measurement (no need for calibration) method of obtaining volume-based size distributions of the particles in a sample. This method passes a laser through the sample and measures the intensity of the (forward) scattered light at different angles (Fig. 1.7a), obtaining what is called a scattering pattern or scattering function. The angle of diffraction of the light is inversely proportional to the particle size, and the intensity of the diffracted beam at any angle is a measure of the number of particles with a specific cross-sectional area in the beam's path (Eshel *et al.*, 2004). The size distribution is obtained from the deconvulsion of the composite scattering pattern using an optimal model, usually either the Fraunhofer or Mie theories of light scattering. The Fraunhofer diffraction model is based on the approximation that the laser beam is parallel and the detector is at a great distance compared with the size of the diffracting particle, and may be used for large particle sizes (particle diameter >20-25 µm approximately).

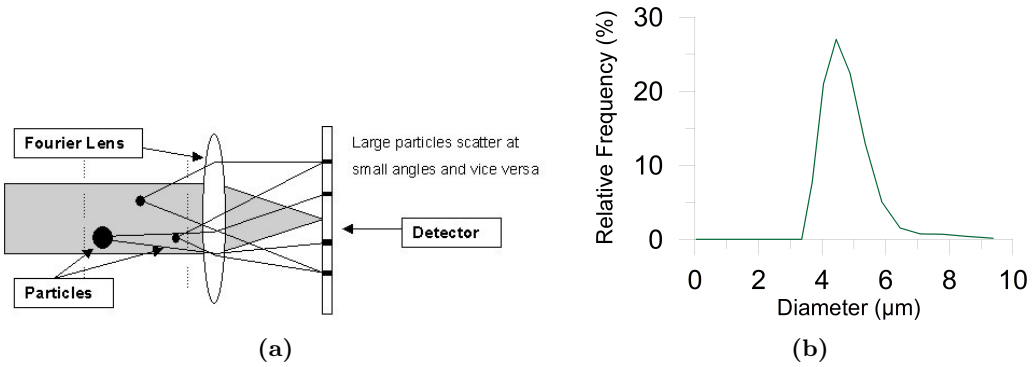
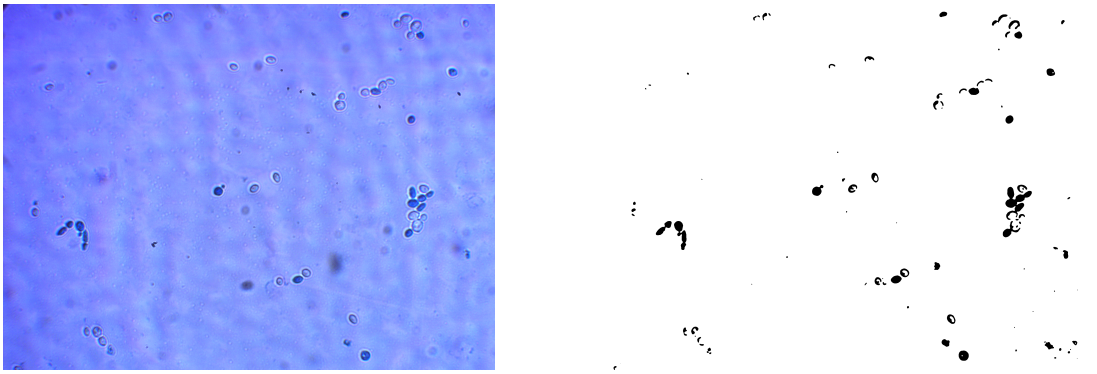


Figure 1.7: Light diffraction principle diagram (a) and a size distribution of aerobically grown stationary phase cells of a *S. cerevisiae* wine yeast strain recorded with a Beckman Coulter LS 13 320 (b). The diagram in (a) is extracted from Beckman Coulter, Inc. (2011). Distribution in (b) corresponds to the same sample from which the distribution of Fig. 1.6b was obtained.

The Mie theory takes into account phenomena other than diffraction and requires knowledge of the refractive index of both the media and the cells being sized. The Mie theory offers an exact solution to the scattering of light from a homogeneous sphere (and not for other particle shapes) and is usable for small and large particles. Light diffraction instruments usually have a high dynamic range (e.g., diameter from 0.017 µm to 2000 µm for the Beckman Coulter LS 13 320). An example size distribution can be seen in Fig. 1.7b.

1.4.4 Image cytometry

Image cytometry represents a wide range of methods to extract quantitative biological information from microscopic images. Essentially, measurements of interest are either stereological (volumes, areas, lengths, profiles, etc.) or photometric (absorbance, fluorescence, luminescence, etc.)(Chieco *et al.*, 2013). Schematically, the steps in image cytometry include the acquisition of (digital) images by microscopy, and the subsequent processing of these images (digital image processing). Most image analysis methods incorporate some form of colorimetric or fluorescent cell staining (e.g., gram staining, diamidino-2-phenylindole -DAPI-, Fluorescent *In Situ* Hybridization or FISH). This provides a means of cell identification or characterization and generation of high contrast images suitable for further processing. Extraction of the desired information from the images usually requires a number of processing steps. These typically includes thresholding, filtering, edge detection, removal of optical artifacts (e.g. fluorescent "halos"), background substraction, pixel averaging, and other transformations (Brehm-Stecher and Johnson, 2004). A simple example is shown in Fig. 1.8. Although not always easy or convenient (Chieco *et al.*, 2013), image processing automatization is also possible. Analysis of video-based images (multiple still images from a time series) can be used to follow individual-cell events with high temporal resolution. Image citometry methods most currently used for measuring cell size in yeasts (see Turner *et al.*, 2012) are: bright field and fluorescent proteins.



(a) **(b)**

Figure 1.8: Digital image cytometry of *S. cerevisiae* var. *bayanus*. An RGB colour image (a) was taken from an aerobical batch culture at 12 h from the beginning of the experiment (x500). The colour image was transformed, after a number of steps, into a binary (black or white) file (b) showing the "objects" of interest. Non-desired objects (small dots, open cells, etc.) can be further avoided before extracting the information wanted (i.e., mean diameter, area, perimeter, etc.) from the binary image.

In the bright field method, the volume of the cell is estimated from area of the focal plane. For *S. cerevisiae*, an ellipsoid shape is often assumed (see, for instance, Tibayrenc *et al.*, 2010). The fluorescent of constitutively expressed fluorophores allows the single-cell protein content to be estimated, giving an indirect measure of the cell size.

1.5 Individual-based modelling

In its core, an Individual-based Model (IbM) explicitly simulates (autonomous) individuals and the behaviour of the population level (i.e., all existing individuals at any given time) emerges from the cumulative behaviour and biotic (among individuals) and abiotic (between the individuals and surrounding medium) interactions at the individual-level. At the same time, the system-level dynamics constrain the behaviour of the individuals. Rules are applied in defining the individuals and the medium behaviour; hence, the descriptor rule-based approach fits the methodology. An IbM can be seen as a particular case of an Agent-based Model (AbM), a term typically used in the field of social sciences. An agent generalizes the concept of the individual and can describe any autonomous entity with its own behaviour and goals (e.g., an enterprise).

Individual-based modelling methodology exemplifies what is called simulation-oriented research, a rapidly growing approach to studying biological systems. It is increasingly being seen as a valuable way to study complex systems, and is used as a bridge between theoretical and experimental approaches. This methodology can be used to partially test new theories when the experimental proof is too erratic, time consuming, tedious or expensive, and in designing and/or delimiting the relevant experiments to be carried out. Individual-based modelling

and simulation is a useful "bottom-up", or "systems", approach where the behaviour of the modelled system emerges from that of the low-level components (see, for instance, Grimm and Railsback, 2005; Hellweger and Bucci, 2009; Railsback and Grimm, 2012; Swinnen *et al.*, 2004). Nonetheless, it is worth stressing that the focus is made on the system (or population) itself and not on the individuals.

1.5.1 Individual-based modelling in microbiology: a definition

Appearing in the in the literature since the 1970s, the individual-based approach is well-established in ecological modelling where most applications have been geared to higher trophic levels. Within the ecological modelling, the IbMs have been defined by Grimm (1999) as "simulation models that treat individuals as unique and discrete entities which have at least one property in addition to age that changes during the life cycle."

Microbes and their properties are inherently more difficult to observe and measure than higher animals. Consequently, microbial ecology is more dependent on technical advances. This fact has undoubtedly contributed to delaying the utilization of individual-based modelling approaches in microbiology. Nevertheless, advances in microbiology and biochemistry have also stimulated an increase in the application of IbMs to microbes (see section 1.3). Some examples of microbial Individual-based Models (μIbMs) are the works of Kreft *et al.* (1998) Dens *et al.* (2005) Bucci *et al.* (2012) Tack *et al.* (2014) and Hellweger *et al.* (2014).

In the microbiological context, and based on the observation that microbe models often do not include age, but size or cell cycle stage, or an explicit representation of the life cycle, Hellweger and Bucci (2009), who review the application of IbMs to microbes, relax this prior definition to "simulation models that treat individuals as unique and discrete entities which have at least two independent properties." Despite the fact that the position may be of capital importance in tackling some microbial systems, the previous definition does not include the position of the individuals in the medium.

1.5.2 Why microbial individual-based models?

The strongest point of the individual-based methodology is its ability to dissect the modelling systems, which finally allow us to gain a deeper understanding of the system itself. From a general point of view, stated arguments for the individual-based approach include the ability to simulate variability among individuals, local interactions, complete life cycles and individual behaviour according to the changing individual internal and external environment (Grimm *et al.*, 2006). Specific motivations of the scientific community in μIbMs mainly include (Hellweger and Bucci, 2009) their ability to resolve intra-population variability (population heterogeneity), their ability to link mechanism at the individual level to the behaviour of the population level (emergence), and the recognition of the inapplicability of the continuum hypothesis.

Population heterogeneity (population structure) is related to the recognition that heterogeneity plays an important role at the system level (see, for instance, Kreft *et al.*, 1998; Malakar

and Barker, 2008; Pin and Baranyi, 2008). In that sense, a strong point for the individual-level approach is that it allows integration of the structure of the population at a given moment (e.g., distribution of the size of the cells of a brewing starting culture) and the following of how this structure evolves dynamically. This facet is and will be strengthened by the stated increasing availability of IbOs. The very existence and availability of IbOs calls for modelling approaches producing outputs that can be compared against this new way of looking at the microbial world, as IbMs do.

The inapplicability of the continuum hypothesis is related to the observation that individuality matters and, due to the high microbial growth rates, can determinate the fate of the population itself, even for large population sizes. For instance, a single enteric bacteria strain can colonize the gut of a single human where about 10^{14} enteric bacteria exist (traveller's diarrhoea). Nevertheless, where the fate of a single (or a few) individual(s) may be crucial is when low inoculum and/or medium heterogeneity come into play. Food contamination, which typically involves contamination by an extremely reduced number of microorganisms (see, Grijspeerdt *et al.*, 2005, and references therein), which are heterogeneous by definition, is a prime example.

The stated availability of IbOs opens new and exciting perspectives for microbiology and individual-based modelling alike. The need for a close collaboration among individual-based modellers and experimentalists has been advocated, either directly or indirectly, by several authors. The combination of individual-based modelling and experimentation has been recently referred by Kreft *et al.* (2013) as microbial Individual-based Ecology (μIbE). μIbE can integrate prior knowledge with new data on the individual and population levels, which, among others things, will directly assist experimentation. However, a key advantage of μIbE holding tremendous potential is the ability to incorporate both lower (i.e., molecular level) and higher levels of organization (communities, ecosystems) into the individual-based framework. Description of the microbial metabolism by embedded intracellular models will drive the individual behaviour more mechanistically, leaving phenomenological descriptions behind. The emergence of a higher trophic level is straightforward for an IbM, whose communities are made up of individual organisms. The interactions of individuals are localized, affecting only those in their immediate vicinity; hence, this is viewed as spatial heterogeneity; and the indirect interactions between individuals of different species (e.g., those mediated by diffusible molecules) are an emergent property, requiring no further *a priori* assumptions or specific parameterization

1.5.3 The individual-based model modelling cycle

An overview of the modelling cycle usually followed in building and using individual-based models will be provided below. This overview aims to allow the reader to better follow the contents of the present thesis. A comprehensible and thorough description of the whole IbM modelling cycle, including model analysis, can be found in Grimm and Railsback (2005) and Railsback and Grimm (2012).

A functional representation of the modelling cycle is shown in Fig. 1.9. It is essential to

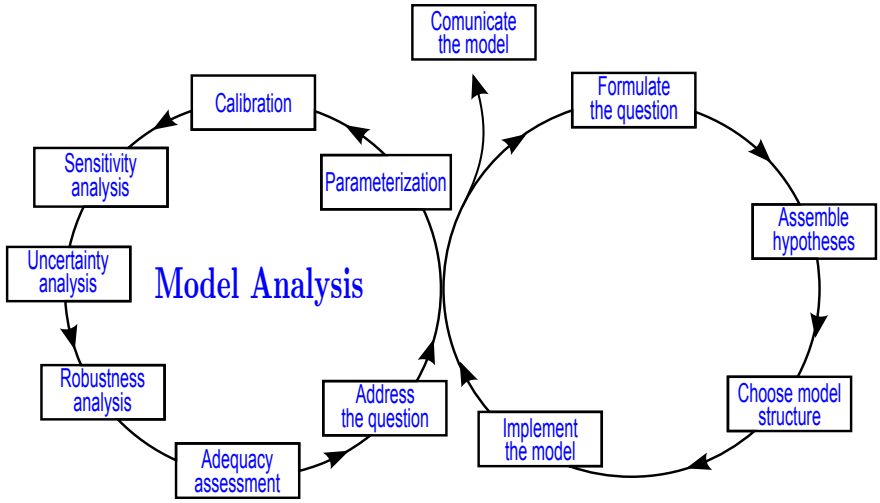


Figure 1.9: The Individual-based Model modelling cycle. This researcher's own elaboration and partial adaptation from Grimm and Railsback (2005).

remember that the modelling is following its cycle iteratively; hence, there is no need for the entire cycle to be completed before a new iteration is undertaken, although it often makes smaller loops between two or more tasks.

Question, hypotheses, and model structure. The modelling cycle starts with a clear research question which will serve as a primary compass and filter for deciding the essential element and processes needed for addressing the question or problem. The elements and processes we considered to be essential, and the establishing of the relationships between them, represents a first conceptual model. The conceptual model is usually formulated verbally and graphically and is based on theory and experience. After the hypotheses have been assembled, the structure of the model has to be chosen. This task leads to deciding on the spatial and temporal extent and degree of detail of the model, its variables and parameters, and the equations and rules to describe the processes identified previously. Variables represent the structure of the model system and parameters, which are used in equations and rules, represent the processes. Producing a written formulation of the model at this stage is highly advisable.

Model implementation. This written formulation (if done) can then be used as a basis to create an animated entity, the simulator, by using mathematics and computer programs. The designing and writing of the computer software for the model is called implementation. Only the simulator can illustrate the consequences of the formulation of the model. This is the most technical part of the modelling cycle and usually requires programming skills, something that may be intimidating for scientists not trained in software design. A process of software verification is (or should be) implicit in the implementation of the model formulation. Additionally, documenting the tests performed in order to provide persuasive evidence of the

correct implementation of the model is best practice when it comes to establishing the reliability of both the model and the modelling work as a whole.

Although each has its own advantages and disavantages, a great many software platforms may be used to implement an IbM. Options range include (Devillers *et al.*, 2010; Grimm and Railsback, 2005), but are not limited to, classic procedural programming languages (e.g., C and Fortran), object-oriented programming languages (e.g., C++ and Java), general high-level modelling environments (e.g., MATLAB, MODSIM, and Simscript), agent-based modelling frameworks and Libraries (e.g. EcoSim, Swarm and RePast), and high-level agent-based modelling environments (e.g., AgentSheets, EcoBeaker, NetLogo, and StarLogo). Very simplistically, lower level platforms require programming from scratch but are computationally more efficient; whereas high level environments add simplified programming specific languages and tools, at the expense of computational efficiency. Computational efficiency within the context of μIbMs, where simulation of a much higher number of individuals than in other areas is typically required, has to be taken into account.

Model analysis. Analysing an IbM means studying the model once it executes so we can understand and improve its performance and use the model to solve the problems it was designed to address. Although different modelling projects may require specific approaches, model analysis can be divided into six main tasks: (i) model parameterization, (ii) model calibration, (iii) sensitivity analysis, (iv) uncertainty analysis, (v) robustness analysis, (vi) assessment of the model adequacy, and (vii) solving the biological problem originally motivating the model development.

The main strategy involved in all these steps is the careful design and realization of *controlled simulation series* (see, for instance, Thiele *et al.*, 2014). In essence, this involves comparing and ranking different versions of the model. The term "different versions" of the model should be interpreted in a broad sense here; meaning, for instance, simulations started with different values for the model parameters, or alternative submodel versions. In order to compare these different model versions, we require a criterion or criteria telling us whether any improvement in the model is discernible or not.

Model parameterization and calibration. Model parameterization and calibration are two key stages. The parameterization stage involves the use of empirical information from real systems to give a range (or a fixed value) to the parameters of the model, enabling us to learn which processes are of paramount importance and which dynamics are most believable. Calibration can be regarded as a particular case of parameterization where we look for accurate values of important parameters by selecting parameter values that cause the model to reproduce experimental observations of the real system. According to the final outcome pursued in this stage, two main ways of model calibration have been defined: categorical and best-fit calibration. Categorical calibration searches for parameter values producing results of the model within a category or range defined as acceptable, while best fit calibration searches for one

unique set of parameter values causing model results which best match some exact criteria, essentially, an optimization. These processes involve (or should involve) comparison of simulated results against (many) experimental results or knowledge at both individual and system level, and ideally, from different time and spatial scales. It is significant that in individual-based methodology, the term calibration is analogous to the parameterization stage (finding values for the model parameters according to certain criteria) of the well-established statistical modelling (see, for instance, Piou *et al.*, 2009).

Sensitivity, uncertainty, and robustness analysis. Sensitivity analysis and uncertainty analysis are special kinds of simulation experiments used to analyse a model in standard, rigorous ways. As highlighted, these use many model runs (i.e. simulation series) to examine the sensitivity of the output of the model to changes in parameter value. The analysis can show how strongly the model represents real-world phenomena, and also helps us understand the relative importance of model processes in the output(s) investigated. This last peculiarity may be used to focus further analysis on a reduced number of model parameters, fixing the value of the less sensitive ones. Uncertainty analysis looks at how uncertainty in parameter values affects the reliability of model results. The modeller specifies probability distributions for selected model inputs of particular importance or with highly uncertain values. The model is executed assigning to the model inputs a value drawn (either randomly or systematically) from the defined distributions. The relative probability of different model results is estimated by repeating the process a sufficient number of times. Robustness analysis is used in different ways in many fields. Grimm and Railsback (2005) use it to describe a general strategy of analysing how robust the outcomes of an IbM are to changes in the inputs or in its assumptions. The focus of the robustness (inverse of sensitivity) is on identifying model results (or better yet, the mechanisms producing them) which are so robust to changes in parameters and model structure that they are likely to be of general significance. This is accomplished by studying a number of similar (simplified versions, unrealistic scenarios, or more complex versions) but distinct models of the same phenomenon.

Assessment of the model adequacy. An important task before the model can be used to solve the original question or problem is to assess the adequacy of the model. This step is considered the "gold" standard and is also referred to as "validation", a term taken directly from the field of software development. The task involves testing independent predictions of the model against experimental data. Strictly speaking, independent predictions refer to patterns or variables not considered in the design and calibration of the model (secondary predictions). Using secondary predictions provides more evidence than using the patterns or variables used to calibrate the model. However, if the same patterns or variables used to calibrate the model are also utilized (a common practice nowadays), it is essential to do it against a dataset not used in the model calibration step. Once the adequacy of the model has been properly verified, *virtual experiments* (i.e., the modelling counterpart of a real experiment) can be used to deal

with the original question (or its reformulation if necessary) from where the modelling cycle started.

Using and communicating the model. Finally, once we have a model in which we are sufficiently confident and that answers our original question or problem, fully or in part, then the model and its results are communicated to the scientific community or to the managers who are going to use it. It is generally assumed that IbMs are more difficult to analyse, understand, and communicate than analytical models. This is due to the fact that IbMs are often described verbally without a clear indication of the equations, algorithms, and schedules that are used in the model (Grimm *et al.*, 2006). Consequently, the outcomes obtained from an IbM are not easily reproduced. This drawback has been greatly mitigated with the proposition (Grimm *et al.*, 2006) and subsequent update (Grimm *et al.*, 2010) of a standard protocol to describe IbMs. Since then, the Overview, Design concepts, and Details (ODD) protocol has become the most currently accepted tool when it comes to facilitating communication and replication of IbMs. Another approach in communicating the model, an one not always possible or convenient, is the inclusion of the whole programming code.

1.5.4 NetLogo and the individual-based model modelling cycle

Over the last few years, a number of dedicated agent-based modelling platforms have become available on the Internet (Devillers *et al.*, 2010). Utilization of a high-level agent based modelling environment offers many advantages. One of the more successful modelling software platforms for agent-based modelling nowadays is NetLogo, which is multiplatform open source software freely downloadable on the net. It already enjoys a significant global presence, and is in continuous development by the Center for Connected Learning and Computer-Based Modelling research group (Wilensky, 1999). It can be safely stated that the main drawback of NetLogo, within the μIbM field, is its low computational efficiency (see section 1.5.3) which does not facilitate the implementation of complicated models with a high number of individuals. Nevertheless, the use of NetLogo for rapid prototyping of model concepts has been encouraged (Grimm and Railsback, 2005). The amount of work required for designing, and implementing the simulator can be greatly reduced by using NetLogo. In addition, the platform provides specialized tools and extensions greatly facilitating the realization and analysis of simulations (e.g., behaviour space, R extension, etc.); and, consequently, the model analysis. Good documentation, existence of code examples, and its talented user community are non-programming benefits derived from this AbM dedicated platform. The ample variety of NetLogo related resources can be consulted on the platform's website (https://ccl.northwestern.edu/netlogo/).

μIbMs have characteristics and essential designs that cannot be described by means of equations and parameters with a synthetic formulation. Although efforts are being made (see section 1.5.3), they do not have a globally accepted common language for their communication, and are usually too broad to be completely described in a single informative or scientific publication. Without the option of having a complete computational code for the simulation

model, its description has tended to be mainly verbal (text) to transmit the information contained in it. It has been suggested that solving the problem of how to communicate IbMs may increase their scientific credibility (Railsback and Grimm, 2012). With NetLogo, the whole programming code of the implemented model is available to the user, and even the model can be saved as applets to be embedded in a web page. The computer code, jointly with a standardized software design and the NetLogo's standard documentation, may completely describe the model, so it can be efficiently communicated. This, in turn, makes a huge contribution to the model's visibility and foments collaboration among scientists.

1.6 The INDISIM framework

The acronym INDISIM stands for INDividual DIScrete SIMulation. It refers to a modelling methodology and also to the software developed in Fortran to examine the outcome of particular models through computer simulations. INDISIM was specifically designed to study bacterial communities and their environment through a mechanistic approach. A short review and selected applications of INDISIM can be found in Ferrer *et al.* (2008). A brief outline of the development of INDISIM methodology is provided here. For a comprehensive reviews of its origins and history, the reader is referred to Prats (2008) and Ferrer (2010).

The origin of this methodology dates back to the 1980s where a first seminal approach was developed. This was further developed by López (1992), and Ginovart (1997), who gradually increased the degree of detail and specification of the models. The modelling methodology was revised, standardized and finally presented by Ginovart *et al.* (2002a), who first coined the definitive name of INDISIM, and explained the methodology in detail. From there, INDISIM evolved through the study of specific cases of interest. INDISIM simulations have succeeded in topics as varied as bacterial growth in agar plates (Ginovart *et al.*, 2002b) and the study of the influence of bacteria size and shape in yoghurt processing (Ginovart *et al.*, 2002c), in which the interaction between two bacterial species (*Streptococcus thermophilus* and *Lactobacillus bulgaricus*) was tackled by means of the study of axenic and mixed cultures.

From there, INDISIM was ready to take on more complex microorganisms, processes, and systems. This marked the birth of several adaptations such as INDISIM-YEAST (Ginovart *et al.*, 2007), which simulates a generic budding yeast; INDISIM-RBC (Ferrer *et al.*, 2007), for studying the spread of the malaria parasite in *in vitro* red blood cell cultures; INDISIM-SOM (Ginovart *et al.*, 2005; Gras, 2004; Gras *et al.*, 2010; 2011), which considers organic matter dynamics and the microbial activity in soils; INDISIM-COMP (Prats *et al.*, 2010), which focuses on the modelling and simulation of the composting process; and INDISIM-*Paracoccus* (Araujo *et al.*, 2014), for studying the activity of the denitrifying bacterium *Paracoccus denitrificans*. A graphical overview of the existing INDISIM simulators is provided in Fig. 1.10. Despite producing valuable knowledge of the biological systems tackled by these simulators, most of the effort has been put into developing the modelling methodology and not in other phases of the cycle (section 1.5.3). Tackling such a casuistic variety with a common methodology provides

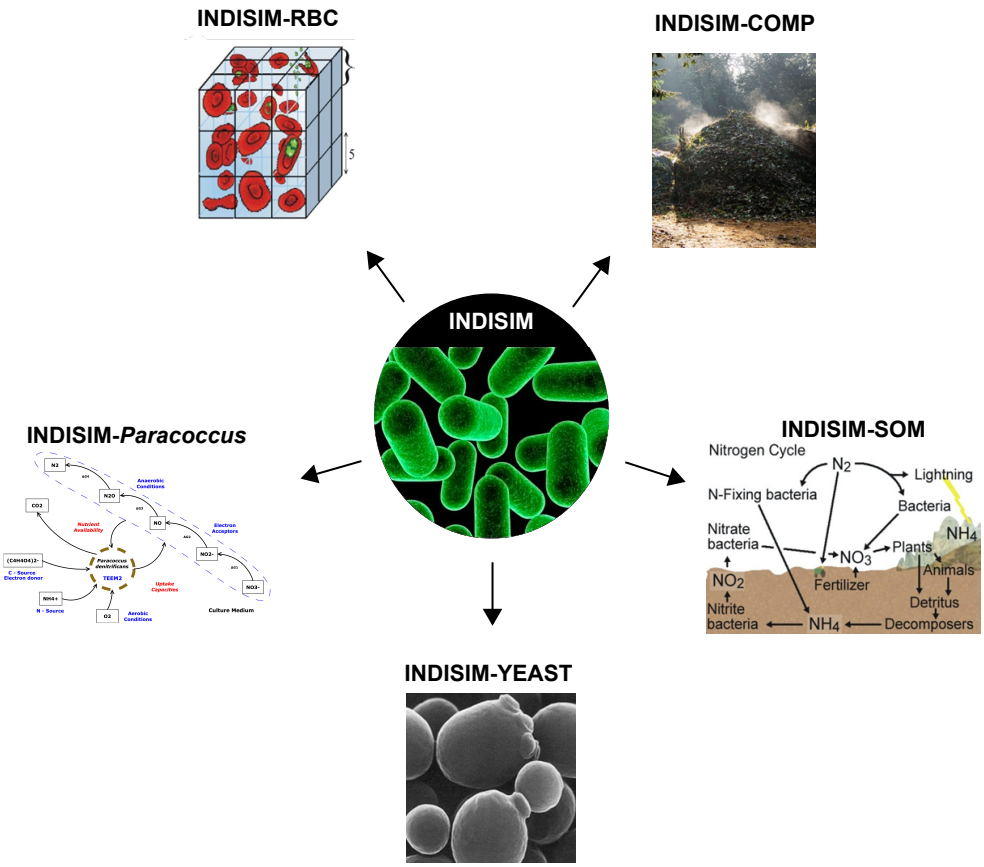


Figure 1.10: INDISIM and INDISIM adaptations existing to date.

a wider perspective of complex systems; allowing development, experience, and a wealth of knowledge to be culled from many areas to be integrated within the most recent models.

1.6.1 INDISIM-YEAST: a general outline

INDISIM-YEAST is the adaptation of INDISIM for tackling the study of the specific characteristics of the yeast cell cycle and its asymmetric budding reproduction. The present section summarizes the INDISIM-YEAST simulator; for a complete description of the methodology the reader is referred to the works of Ginovart *et al.* (2002a), and Ginovart *et al.* (2007). An interactive demonstration version of INDISIM-YEAST (Ginovart and Cañadas, 2008) is accessible from https://aneto.upc.es/simulacio/hoja-portada.html.

The spatially explicit IbM INDISIM-YEAST is discrete in space and time. The modelled space is split into a set of regular divisions, the spatial cells, while the modelled events occur at finite and regular intervals, the computer or time steps. The simulation length is defined by the number of time steps the simulation is allowed to continue for. Two main entities are taken into account within the model: the yeast cell and the spatial cell. The population is the ensemble of microbial cells and the environment is the set of spatial cells. The modelled microbial individual represent a generic budding yeast with a pure fermentative metabolism, that is, taking up nutrient particles (glucose) and excreting a main metabolite product (ethanol). The state of a population composed of N microbial cells at a given time step t is described with the matrix $P_N(t)$.

$$P_N(t) = \left\{ E_i \left[e_1^i(t),\ e_2^i(t),\ ...,\ e_9^i(t) \right] \right\}_{i=1,2,\ ...,\ N}$$

The state of each yeast cell is defined with a vector $E_i(t)$ containing its individual characteristics and variables: $e_1^i(t)$, and $e_2^i(t)$ identify its position in the spatial domain; $e_3^i(t)$, its biomass, which is related by the model to spherical geometry in order to evaluate its cellular surface; $e_4^i(t)$, its genealogical age as the number of bud scars on the cellular membrane; $e_5^i(t)$, the reproduction phase in the cellular cycle in which the cell currently is, namely the unbudded or budding phase, i.e. Phase 1-the unbudded phase, when the cell gets ready to create a new cell (the bud), and Phase 2-the budding phase, in which the daughter cell-genealogical age 0 (virgin cell) grows until it separates from the parent cell, leaving behind another scar; $e_6^i(t)$, its "start mass", the mass required to change from the unbudded to budding phase; $e_7^i(t)$, the minimum growth of its biomass for the budding phase; $e_8^i(t)$, the minimum time required to complete the budding phase; $e_9^i(t)$, its survival time without satisfying its metabolic requirements. The components of $E_i(t)$, the individual variables, may change throughout the simulation according to the set of rules governing individuals.

The simulated area is a regular grid (2D), with liquid medium and yeast, which is subject to appropriate boundary conditions. The modelled space is divided into spatial cells of coordinates $[x, y]$. The state of the environment at a given time step t is described with the matrix $D(t)$.

$$D\,(t) = \{S_{xy}\,[s_1^{xy}\,(t)\,,\ s_2^{xy}\,(t)]\}_{x=1,\ \ldots,\ k;\ y=1,\ \ldots,\ k}$$

k stands for the number of spatial cells per dimension. Each spatial cell represents the local environment of the yeast cells, and is described with the vector $S_{xy}(t)$. The components of this vector denote: $s_1^{xy}(t)$, the number of nutrient (glucose) particles, and, $s_2^{xy}(t)$, the number of metabolite particles (ethanol). Apart from the activity of the individuals, which also modify their local environment, the processes that affect each spatial cell are described by the set of rules governing the environment (e.g. substrate diffusion, among others).

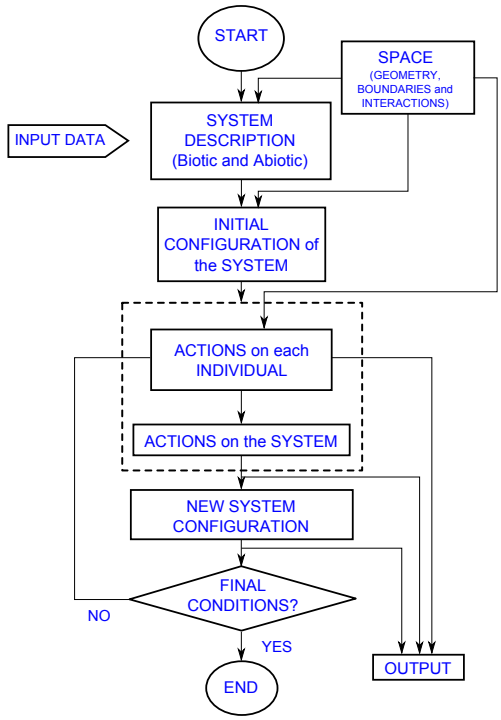


Figure 1.11: General basic flow chart of INDISIM-YEAST.

The operating procedure of the simulator INDISIM-YEAST (Fig. 1.11) can be partitioned in different sections: (i) initialization of the system, with the entrance of input data that fixes the initial configuration of the whole system; (ii) the main loop (time step), where the rules governing the system are applied recurrently until the simulation ends, and where new configurations of the system originate at each time step, and (iii) data output at the end of each time step or at the end of the simulation to obtain the results. The rules governing the whole system, applied at each time step, can be further divided into the actions on the individuals and actions on the (abiotic) medium. Individual actions taken into account are motion, uptake, metabolism, reproduction, and viability (Fig. 1.12). Possible actions on the environment to simulate different culture strategies are depicted in Fig. 1.13.

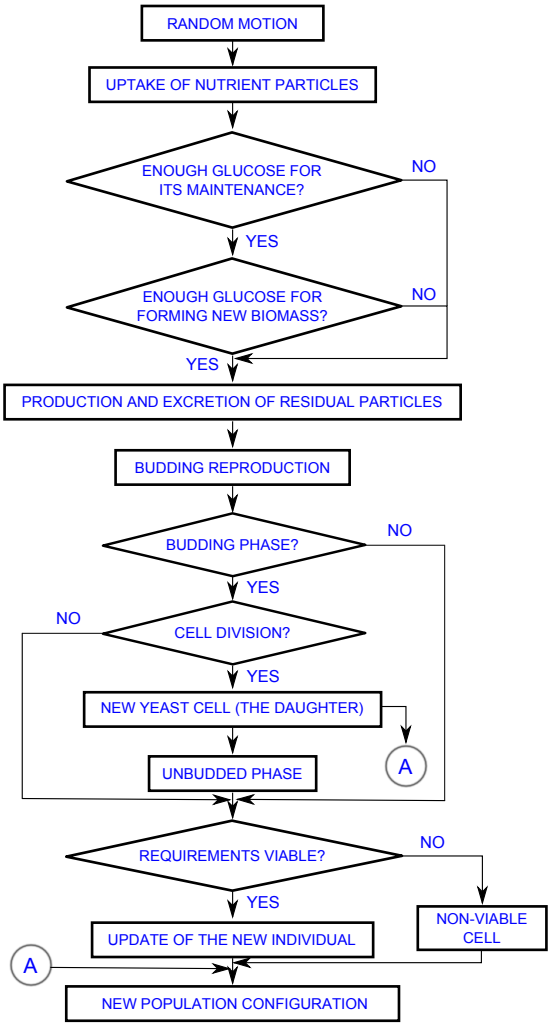


Figure 1.12: Actions implemented in the simulator INDISIM-YEAST followed at each time step by the simulated yeast.

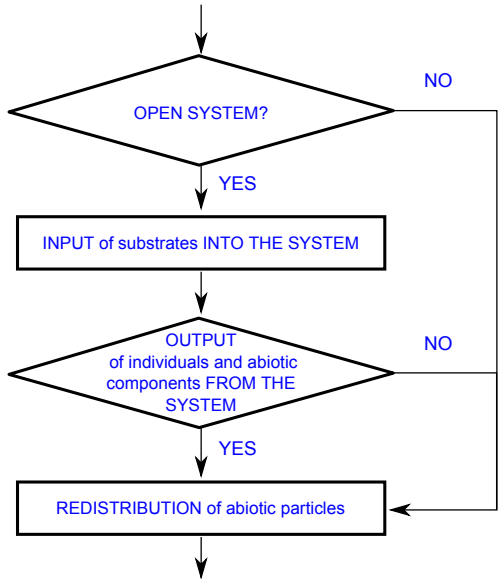


Figure 1.13: Possible alternative actions that may be applied to the simulated abiotic medium of INDISIM-YEAST. The actions applied have to be set accordingly in order to simulate different experimental culture strategies.

Processes taken into account in the model are described and controlled by mathematical expressions and rules driven by a set of time-dependent variables. The yeast model implemented in the simulator takes into account the effects arising from: the bud scars, as these affect the cellular membrane; the excess of nutrient concentration, as it may also induce inhibition in its uptake; and the growth arrest as a consequence of the metabolic final product. Random variables and/or random numbers are used to characterize the individual yeast cell and the individual actions for the update of the set of rules.

Rules governing the individuals and their local environment are implemented as independent sub-models that can be switched on and off at will. This compartmentalization allows building models with increasing complexity and the comparison of the outcome of different versions of the model with each other. Such a procedure provides the staggered study of complex systems and allows the appropriate degree of complexity of the models to be determined.

INDISIM-YEAST can track the evolution of single individuals and single spatial cells. It can also store data at any moment during the time step, thus following in detail the course of an event within the time step. This feature allows the comprehensive monitoring of any process occurring in the model. In other words, every piece of the virtual world created with INDISIM can be torn apart. This allows comprehension of the *causality* leading to the model outputs and, lastly, gaining a deeper understanding of the system being studied.

1.6.2 INDISIM-YEAST: weaknesses

INDISIM-YEAST allows a number of interesting studies to be undertaken. However, a number of its characteristics can be seen as strengths in certain situations and weaknesses in others, thus limiting its applicability. These weaknesses are summarized below.

INDISIM-YEAST, as seen in the previous section, uses a bi-dimensional lattice of discrete spatial cells. This allows computer power to be saved but may not biotechnical applications to be tackled where the spatial occupation of a volume is relevant to the behaviour of the system, as, for instance, in yeast sedimentation processes and/or in non-homogeneous bioreactors, which are closer to reality.

Conceptually, the yeast biology (uptake, biomass composition, metabolic alternatives, internal nutrient pools, etc.) modelled in INDISIM-YEAST is extremely simple. The model was originally designed to study the implications on the system of an asymmetrical division scheme. The uptaken nutrient particles are catabolized to cover maintenance of the cell and, if in excess, to produce new biomass with a certain yield. Therefore, the nutrient acts as an energy and carbon source. It is significant that no N source is needed (biomass only made of C). Since an alcoholic fermentation is modelled, the catabolism of the nutrient produces the excretion to the medium of ethanol as the main metabolite. A number of weaknesses of INDISIM-YEAST are related to this conceptualization of the yeast metabolism. A common observation regarding important industrial yeast cultures is the limitation of growth by N sources (Ribéreau-Gayon *et al.*, 1998; Walker, 1998). A simulator solely taking into account glucose as a nutrient particle is built under the assumption that N sources are excessive and thus do not limit yeast growth, making the reproduction of this experimental observation intractable. Exclusion of the aerobic metabolism serially limits the applicability of the model to simulate the high number of experimental situations that can experience yeast in many biotechnological applications.

Nevertheless, the most limiting INDISIM-YEAST weakness is probably its inability to produce directly comparable outputs. INDISIM-YEAST uses arbitrary units, the so-called simulation units, which permit the realization of meaningful qualitative studies. Qualitative simulations may be a valuable tool for the modeller to understand and to compare the relative behaviour of the system under varying scenarios. Unluckily, quantitative experimentally falsifiable studies, a *sine qua non* for using the model for predictive purposes, are simply unachievable.

Chapter 2

Objectives of the thesis

The aim of the present dissertation is to advance the development of an IbM methodology to tackle the study of microbial systems driven by the relevant yeast *S. cerevisiae*. This general aim is addressed by establishing four specific objectives:

1. To assess the adequacy of the existing IbM model INDISIM-YEAST, based on a generic budding yeast, to be used as a core model to develop a simulator based on *S. cerevisiae*.

2. To verify and asses the diversity in experimental individually-oriented observations of *S. cerevisiae* under different growth conditions and in different stages of the growth curve.

3. To develop, parameterize, and calibrate an individual-based simulator to tackle the study of microbial systems driven by *S. cerevisiae* taking into account the metabolic alternatives of this yeast.

4. To communicate efficiently, increase accessibility, and favour usability of the developed methodology to both skilled and unskilled potential users.

Chapter 3

Assessment of the INDISIM-YEAST adequacy as a core model

3.1 Introduction

The generic yeast model INDISIM-YEAST was mainly designed to deal with the implications of the asymmetric division on the growth of the population. Nevertheless, as previously stated within this document, most of the work with INDISIM-YEAST focused on the first part of the modelling cycle, that is, the design and implementation of the model (right circle of the figure 1.9 on page 16). The model was not thoroughly analysed or used in any specific application. Taking into account that one of the points of this thesis aimed, using INDISIM-YEAST as a core simulator, to build a completely new and improved model focused on *S. cerevisiae*, testing the adequacy of this qualitative model, i.e. evaluating whether the model captures the essentials of the modelled yeast system, has been deemed essential.

The lack of INDISIM-YEAST analysis is addressed in this chapter with the realization, and subsequent analysis, of controlled simulation series to investigate the effects of specific characteristics of the initial inocula — the inoculum size (i.e., initial cells/ml); and the genealogical age (number of cell divisions experienced) — on the, on the dynamics of a virtual yeast culture in liquid medium during the first stages of a batch culture (i.e., lag and exponential phase). Both inoculum size and the genealogical age of the budding yeasts have been recognised as being important factors influencing industrial yeast fermentations. In particular, in the brewing industry, which is unique in the sense that it reuses the yeast a number of times (a process called "serial repitching"), the cropping practices may result in differing initial distribution of genealogical ages, so that immediate and long-term fermentation performance is conditioned by the characteristics of these reused inocula. This has been taken into account in designing

the simulation series. The analysis was published in two papers which now form the sections 3.2 (Ginovart *et al.*, 2011b) and 3.3 (Ginovart *et al.*, 2011a) of the present chapter.

3.2 Exploring the lag phase and growth initiation of a yeast culture by means of an individual-based model

Publication details

Authors:	Marta Ginovart, Clara Prats, **Xavier Portell**, and Moises Silbert
Title:	Exploring the lag phase and growth initiation of a yeast culture by means of an individual-based model
Ref.:	*Food Microbiol.* 28.4, 810-817. 2011
DOI:	10.1016/j.fm.2010.05.004
I. Fac.:	3.283*

* Journal Citation Reports 2011

ATENCIÓ ¡
Les pàgines 32 a 39 de la tesi contenen l'article, que es pot consultar a la web de l'editor

ATENCIÓN ¡

Las páginas 32 a 39 de la tesis contienen el artículo, que puede consultarse en el web del editor

ATTENTION ¡

Pages 32 to 39 of the thesis are availables at the editor's web

http://www.sciencedirect.com/science/article/pii/S074000201000095X

3.3 Analysis of the effect of inoculum characteristics on the first stages of a growing yeast population in beer fermentations by means of an individual-based model

Publication details

Authors:	Marta Ginovart, Clara Prats, **Xavier Portell**, and Moisés Silbert
Title:	Analysis of the effect of inoculum characteristics on the first stages of a growing yeast population in beer fermentations by means of an individual-based Model
Ref.:	*J. Ind. Microbiol. Biotechnol.* 38.1, 153-165. 2011
DOI:	10.1007/s10295-010-0840-4
I. Fac.:	2.735*

* Journal Citation Reports 2011

ATENCIÓ ¡
Les pàgines 41 a 53 de la tesi contenen l'article, que es pot consultar a la web de l'editor

ATENCIÓN ¡

Las páginas 41 a 53 de la tesis contienen el artículo, que puede consultarse en el web del editor

ATTENTION ¡

Pages 41 to 53 of the thesis are availables at the editor's web

http://link.springer.com/article/10.1007%2Fs10295-010-0840-4

Chapter 4

Saccharomyces cerevisiae size distribution under aerobic and microaerophilic conditions

4.1 Introduction

Experimental knowledge is of interest in and of itself. Not just from an experimental standpoint, a quantitative IbM frequently also requires extensive experimental data that can be used in a number of model analysis steps (see section 1.5.3). Ideally, both population-level and individual-level (or IbO) data have to be used in order to increase confidence in the model reliability. Experimental IbOs are usually summarized as frequency distributions, which illustrate the state of the single individuals (or population structure) at a particular time. IbMs produce and are naturally able to integrate this information, predicting the future development of the system. One of the variables which is most biologically significant and directly able to be used in a model is the size of the individuals (see section 1.4 on page 9). Because of the tight coupling between cell growth and division this information is necessary in assessing the performance of the reproduction submodel of an IbM of *S. cerevisiae*.

As stated in section 1.2.1, *S. cerevisiae* is able to use glucose differently depending on the prevailing medium conditions, with important implications for the process in question. Glucose utilization, along with the complexity of the different processes that take place in each phase of the yeast growth (lag phase, exponential phase, stationary phase), impact on the evolution of, among other things, the cell size distribution.

S. cerevisiae batch cultures evolving under different oxygen concentrations were grown at the **Parc Cientific de Barcelona** (Universitat de Barcelona) and analysed using three experimental techniques: flow cytometry (section 1.4.1), electric particle analysis (section 1.4.2), and light diffraction (section 1.4.3). At the beginning of this thesis, the raw data from the

stated experiments was in the possession of the MOSIMBIO research group. These raw data were an interesting input for further modelling steps but needed major preparation and synthesis. The process culminated with the publication of the research papers included below in sections 4.2 (Portell *et al.*, 2011b) and 4.3 (Portell *et al.*, 2011a).

4.2 Population analysis of a commercial *Saccharomyces cerevisiae* wine yeast in batch culture by electric particle analysis, light diffraction, and flow cytometry

Publication details

Authors:	**Xavier Portell**, Marta Ginovart, Rosa Carbó, Anna Gras, and Josep Vives-Rego
Title:	Population analysis of a commercial *Saccharomyces cerevisiae* wine yeast in a batch culture by electric particle analysis, light diffraction and flow cytometry
Ref.:	*FEMS Yeast Res.* 11.1, 18-28. 2011
DOI:	10.1111/j.1567-1364.2010.00682.x
I. Fac.:	2.403*

* Journal Citation Reports 2011

ATENCIÓ ¡
Les pàgines 58 a 68 de la tesi contenen l'article, que es pot consultar a la web de l'editor

ATENCIÓN ¡

Las páginas 58 a 68 de la tesis contienen el artículo, que puede consultarse en el web del editor

ATTENTION ¡

Pages 58 to 68 of the thesis are availables at the editor's web

http://onlinelibrary.wiley.com/doi/10.1111/j.1567-1364.2010.00682.x/full

4.3 Differences in stationary-phase cells of a commercial *Saccharomyces cerevisiae* wine yeast grown in aerobic and microaerophilic batch cultures assessed by electric particle analysis, light diffraction, and flow cytometry

Publication details

Authors:	**Xavier Portell**, Marta Ginovart, Rosa Carbó, and Josep Vives-Rego
Title:	Differences in stationary-phase cells of a commercial *Saccharomyces cerevisiae* wine yeast grown in aerobic and microaerophilic batch cultures assessed by electric particle analysis, light diffraction and flow cytometry
Ref.:	*J. Ind. Microbiol. Biotechnol.* 38.1, 141-151. 2011
DOI:	10.1007/s10295-010-0839-x
I. Fac.:	2.735*

* Journal Citation Reports 2011

ATENCIÓ ¡
Les pàgines 70 a 80 de la tesi contenen l'article, que es pot consultar a la web de l'editors

ATENCIÓN ¡

Las páginas 70 a 80 de la tesis contienen el artículo, que puede consultarse en el web del editor

ATTENTION ¡

Pages 70 to 80 of the thesis are availables at the editor's web

http://link.springer.com/article/10.1007%2Fs10295-010-0839-x

Chapter 5

INDISIM-*Saccha*: an individual-based model of *Saccharomyces cerevisiae*

5.1 Introduction

Taking the generic qualitative yeast model INDISIM-YEAST as a core model, a quantitative model focused on the *S. cerevisiae* biology has been built. This model has been termed INDISIM-*Saccha*, already introduced to the scientific community by the work of Portell *et al.* (2014b), and shown in section 5.2 of the present manuscript. Taking into account the great importance of the fermentative metabolism, which leads to ethanol formation, in classical and still dominant (at least mass-wise) biotechnologies worldwide, this first contribution focused on fermentative processes. An adaptation of INDISIM-*Saccha* to deal with aerobic *S. cerevisiae* growth, the dominant and desirable metabolism for, amongst others, biomass production (e.g., baker's yeast production), was undertaken subsequently. This adaptation was introduced to the scientific community by the work of Portell *et al.* (2014a) [1], and also shown below in section 5.3.

Supplementary Material published online by Portell *et al.* (2014b) can be found in appendixes A and B, while Supplementary Material by Portell *et al.* (2014a) are found in the appendix C.

[1] submited to Ecological Modelling

5.2 INDISIM-*Saccha*, an individual-based model to tackle *Saccharomyces cerevisiae* fermentations

Publication details

Authors:	**Xavier Portell**, Anna Gras, and Marta Ginovart
Title:	INDISIM-*Saccha*, an individual-based model to tackle *Saccharomyces cerevisiae* fermentations
Ref.:	*Ecol. Model.* 279, 12-23. 2014
DOI:	10.1016/j.ecolmodel.2014.02.007
I. Fac.:	2.326*

* Journal Citation Reports 2013

ATENCIÓ ¡
Les pàgines 83 a 94 de la tesi contenen l'article, que es pot consultar a la web de l'editor

ATENCIÓN ¡

Las páginas 83 a 94 de la tesis contienen el artículo, que puede consultarse en el web del editor

ATTENTION ¡

Pages 83 to 94 of the thesis are availables at the editor's web

http://www.sciencedirect.com/science/article/pii/S0304380014000921

5.3 Expanding INDISIM-*Saccha* metabolism to tackle aerobic *Saccharomyces cerevisiae* cultures

Publication details

Authors:	**Xavier Portell**, Anna Gras, Rosa Carbó, and Marta Ginovart
Title:	Expanding INDISIM-*Saccha* metabolism to tackle aerobic *Saccharomyces cerevisiae* cultures
Ref.:	*Ecol. Model.* Submitted. 2014

Expanding INDISIM-*Saccha* metabolism to tackle aerobic *Saccharomyces cerevisiae* cultures

Xavier Portell[a,*], Anna Gras[a], Rosa Carbó[a], Marta Ginovart[b]

[a]*Department of Agri-Food Engineering and Biotechnology, Escola Superior d'Agricultura de Barcelona, Universitat Politècnica de Catalunya, Campus Baix Llobregat, Esteve Terradas 8, 08860 Castelldefels, Barcelona, Spain.*
[b]*Department of Applied Mathematics III, Escola Superior d'Agricultura de Barcelona, Universitat Politècnica de Catalunya, Campus Baix Llobregat, Esteve Terradas 8, 08860 Castelldefels, Barcelona, Spain.*

Abstract

Saccharomyces cerevisiae is one of the yeasts with a major economic, social and health significance in human culture. Depending on the growth conditions being experienced by the cell, the *S. cerevisiae* growth can proceed via fermentative, respirative, or respirofermentative (a.k.a, oxido-reductive) metabolism. An individual-based model (IBM) of the yeast *Saccharomyces cerevisiae* was developed, parameterized, calibrated, and used to examine *in silico* changes in the ethanol production kinetics resulting from differences in the size distribution of the yeast cells that make up the inoculum, and recently published (Portell *et al.* 2014, *Ecol. Model.* 279: 12-23). This model, which incorporated a fermentative (anaerobic) yeast metabolism, was named as INDISIM-*Saccha*. The contribution in question extends the existing IBM model INDISIM-*Saccha* to analyse the dynamics of *S. cerevisiae* cultures evolving in a liquid medium with oxygen. The model description is undertaken following the ODD (Overview, Design concepts, and Details) protocol for describing individual- and agent-based models. The performance of the approach is evaluated by comparing values of the glucose, ethanol and cell density evolutions against their experimental homologous points collected in two experimental conditions differing in oxygen level. The simulated results shown in this contribution suggest that the approach presented holds enough potential to reproduce and investigate aerobic batch cultures of *S. cerevisiae*.

Keywords: Yeast, *Saccharomyces cerevisiae*, aerobic, microaerophile, batch culture, individual-based modelling

1. Introduction

Saccharomyces cerevisiae is one of the yeasts with a major economic, social and health significance in human culture. It has been the subject of myriad basic and applied studies, with an increasing number nowadays being centred on the individual level (e.g., Hellweger et al., 2014; Bermejo et al., 2011; Urban et al., 2011; Gustavsson et al., 2012). Depending on the growth conditions experienced by the cell, the *S. cerevisiae* growth can proceed via fermentative, respirative, or respirofermentative (a.k.a, oxido-reductive) metabolism. Given the variety of sugar

*Corresponding author. Tel.: +34 935521136.
Email addresses: xavier.portell@upc.edu (Xavier Portell), anna.gras@upc.edu (Anna Gras), rosa.carbo@upc.edu (Rosa Carbó), marta.ginovart@upc.edu (Marta Ginovart)

catabolic alternatives, the complexity of their regulation, and an incomplete knowledge of the precise mechanisms controlling them (Alexander and Jeffries, 1990; Walker, 1998), it is difficult to build a conceptual model suitable for use throughout the complete range of growth conditions involving this yeast.

A recently published (Portell et al., 2014) Individual-based Model (IBM) of the yeast *S. cerevisiae* was developed, parameterized, calibrated, and used to examine *in silico* changes in the ethanol production kinetics resulting from differences in the size distribution of the yeast cells that make up the inoculum. This model, which incorporated a fermentative (anaerobic) yeast metabolism, was named INDISIM-*Saccha*. Due to the specific properties of the individual-based approach, an IBM accounting for all the *S. cerevisiae* sugar catabolic alternatives, which are mainly regulated by the availability of glucose and oxygen, would greatly improve our knowledge within changing or heterogeneous media and/or close to the regulatory boundary limits.

The contribution in question aims to test whether the INDISIM-*Saccha* methodology extended by incorporating an oxygen-using metabolic alternative is also capable of tackling experimental cultures with varying oxygen concentrations.

2. Material and methods

2.1. Experimental data

Data from the *S. cerevisiae* strain DV10 grown in two growth conditions, aerobic and microaerophilic, were obtained. The medium used in the aerobic growth tests was composed of 10 g/l glucose, 3 g/l yeast extract, and 3 g/l of casein peptone. In microaerophilic conditions, the medium was supplemented with 0.5 g/l sodium thioglycolate and 0.001 g/l resazurin. The medium pH was adjusted to 3.5 with ortophosforic acid.

Inocula for the experiments were grown at 27 °C for 72 hours in the same media used later (i.e., supplementing microaerophilic preculture with sodium thioglycolate and resazurin). 250 ml flasks covered with sterile cotton and filled with 100 ml of the corresponding medium, also subject to magnetic stirring (300 rpm), were used for aerobic preculture conditions. 50 ml tubes fitted with a screw cap and filled with 50 ml of the corresponding medium were used for microaerophilic preculture conditions.

Growth kinetics in aerobic and microaerophilic conditions were conducted at 27 °C in 1 l flasks filled with 600 ml of broth, and magnetically stirred (300 rpm). Viable population, expressed as colony forming units, and the glucose concentration were determined regularly throughout the 30 hours of the study. Dissolved oxygen concentration was determined at the beginning of the culture while ethanol concentration was determined at 18 hours and at the end of the study.

Colony forming units were determined by using the pour plate method. Glucose concentration was measured by high-performance liquid chromatography (HPLC; Bekman). Ethanol concentration was determined with a Hewlett Packard 5890 Series II GC equipped with flame ionization detection using nitrogen as carrier gas. Initial dissolved oxygen concentration was determined with an oxygen electrode (OxyGuard, Handy Polaris).

2.2. Model description

The description of the model has been undertaken following the ODD protocol for describing individual- and agent-based models (Grimm et al., 2006, 2010). Nevertheless, for practical

reasons, we only provide here a minimal basic factual description of the model in order to establish the context for the reader. For a complete description of the model, the reader is referred to the ODD version found in the Supplementary Materials and in the publications section of the web site: `http://mosimbio.upc.edu`. Although many model parts shown have been described elsewhere (Portell et al., 2014, see Supplementary Materials), the ODD fully describes the model used in order to make a self-contained contribution. In line with the work of Scheller *et al*. (Scheller et al., 2009), which advocates increasing the reliability of ecological models, the ODD provided also includes a section about the testing of the model implemented.

New features of the model described here are: (i) introduction of the oxygen as a metabolic substrate for the yeast; (ii) utilization of aerobic or anaerobic catabolic pathways according to the local environment of the individual (see section 2.2.2); (iii) introduction of an individual adaptation time (individual lag, see section 2.2.3); and (iv) actions on the medium adapted to the experimental conditions faced by the model in this contribution (stirring, oxygen input, and oxygen reduction by external agents; see section 2.2.4).

2.2.1. Model overview

The model has been developed to analyse the dynamics of *S. cerevisiae* batch cultures evolving in a stirred liquid medium with oxygen and under culture conditions not promoting a noticeable Crabtree effect (i.e., low initial glucose - below about 10 g/l).

The model takes into account three entities: *S. cerevisiae* individual cells, spatial cells, and the environment. The simulated space is a cube which is divided into spatial cells of equal size and shape. The environment simulates a liquid medium enclosed in a cube whose faces do not allow the ingress or egress of either organic or inorganic elements, with the exception of molecular oxygen (see section 2.2.4). The temporal evolution of the system is divided into equal intervals associated with time steps.

At each time step existing yeast cells perform, sequentially, the following set of actions: uptake, metabolism, reproduction, and lifespan. Yeast activity is locally described; therefore actions of individuals take place and only affect the spatial cell they are in. At any given time step, once all existing individuals have acted, the actions on the medium, which have to be set according to the experimental practice to be simulated and account for the molecular oxygen input, are performed.

In the model, four substrates can be taken up by the yeast cells: glucose (GLU), organic N (CN), ammonium (NH_4), and molecular oxygen (O_2). CN is assumed to have the following elemental formula: $C_{a_C}N$, being, thus, a C and N source simultaneously.

Yeast cells experience a lag period; a time period where they are internally adapting in order to be able to experience cellular growth. The metabolism submodel assumes the respirative catabolic pathway (glycolysis and Krebs cycle) as the first option in achieving metabolic energy. Nevertheless, it is assumed that the cell is also able to use the fermentative catabolic pathway (glycolysis and alcoholic fermentation) once the uptaken O_2 is depleted. The uptaken glucose is used to cover the yeast maintenance requirements of the cell. Once the maintenance energy is covered, the cell is able to experience growth (i.e. new mass synthesis) for which it needs a C source, an N source, and a specific amount of energy. The yeast accumulates reserve carbohydrates when the glucose content in the spatial cell is below a certain threshold. The reproduction submodel assumes that the cellular cycle involves two differentiated phases: the unbudded and the budded phase. The unbudded phase covers most of phase G1 and a very small fraction of phase S in the traditional division of the eukaryote cell cycle (see, for instance, Walker, 1998);

while the budded phase covers a small fraction of G1, most of S and all of G2 and M. Chronological lifespan (the consequence of cumulative irreversible damage to intracellular components) and replicative lifespan (the consequence of the number of cell divisions experienced) are taken into account within the lifespan submodel. Yeast lysis is not modelled; therefore, no return of biomass components to the medium is modelled.

2.2.2. Respirative metabolism submodel

The model assumes the respirative catabolic pathway (glycolysis and Krebs cycle, Eq. 1) as the first option in achieving metabolic energy. Nevertheless, since respiration is totally optional for *S. cerevisiae* (Alexander and Jeffries, 1990), it is assumed that the cell is also capable of using the fermentative catabolic pathway (glycolysis and alcoholic fermentation, Eq. 2) once the uptaken O_2 is depleted.

$$C_6H_{12}O_6 + 6\ O_2 \rightarrow 6\ H_2O + 6\ CO_2 + 36\ ATP \tag{1}$$

$$C_6H_{12}O_6 \rightarrow 2\ C_2H_6O + 2\ CO_2 + 2\ ATP \tag{2}$$

A yeast cell requires energy for its cellular maintenance, and it is assumed that this energy is proportional to its structural mass ($M(t)$), and affected by the respirative maintenance rate parameter (e_R), and the local content of ethanol through the parameter p_{OH}. If the uptaken O_2 is depleted and a fermentative metabolism is required to cover maintenance, the non-covered respirative maintenance energy is changed to fermentative requirements via the relationship between the energy obtained from both metabolic pathways ($\epsilon_{R/F}$). In case the uptaken amount of glucose is not enough to cover maintenance, the cell uses its existing reserve carbohydrates. In this case, it is assumed that the cell is not able to create new biomass in the current time step. The submodel takes into account that the maintenance requirements of a yeast cell have to be fulfilled before it is able to grow or accumulate new reserve carbohydrates.

The modelled yeast creates reserve carbohydrates when the glucose content in the spatial cell is below a threshold (l_S), otherwise it can synthesize new mass. To create biomass, a C source, an N source and a specific amount of energy are required, which may be acquired through respirative or fermentative metabolism with different biomass yields. The respirative metabolism is taken into account first and succeeded by the fermentative alternative if O_2 becomes limiting. As a consequence of the mass creation, CO_2 and ethanol are released into the local medium.

2.2.3. Individual lag modelling

Experimental conditions described above fostered a population lag phase which appeared once the inocula started growing in the new conditions. This prompted the introduction of the individual lag phase parameter, something unnecessary in the work of Portell *et al.* (Portell et al., 2014).

The population lag phase is a period of zero growth where the cells adapt themselves before division is possible. A population lag may result from a variety of microscopic causes (Prats et al., 2006) and mainly occurs when inoculated cells experience a change in nutritional status, alterations in physical growth conditions, or when cells need to recover from damage, which is typical of previous stressing conditions (e.g., stationary phase). This last case has been assumed and modelled in a simple way. Inoculum cells are assumed to have an adaptation period or individual lag time (T_a) that has to be overcame before starting cell growth. Until T_a

is achieved, only maintenance requirements are taken into account and, consequently, the cell cannot increase its own mass or accumulate internal reserves.

2.2.4. Actions on the medium

The model simulates a stirred batch culture which may or may not be supplemented with sodium thioglycolate, commonly used as a reducing agent for molecular oxygen in laboratory practices. The action of sodium tioglycolate has been modelled by means of an initial sodium thioglycolate pool which is assumed to reduce a fraction f_{O2} of the inputted O_2 in the medium at the current time step. Therefore, both the sodium thioglycolate and oxygen pools decrease.

At every time step, as a consequence of stirring, the position of the individuals is changed randomly within the spatial domain, the atmospheric oxygen enters the medium, and soluble substances (glucose, ethanol, ammonium, organic N, and oxygen) are redistributed uniformly throughout the spatial domain. A proportion v_{O2} of the difference between the solubilized O_2 at equilibrium and the current O_2 concentrations is assumed to input to the medium at every time step.

3. Results and discussion

In order to assess the implemented metabolism performance to reproduce the stirred laboratory experiments of *S. cerevisiae* with oxygen, the glucose, ethanol and cell density determinations collected in two experimental conditions, aerobic and microaerophilic, were compared against simulated data.

The detailed initialization mechanism of the simulations can be seen in the Supplementary Material (Section A.5), while the values adopted are shown in Table 1. Since no experimental measure was available, setting the initial concentration of the N species was done according to the theoretical N contents from the medium ingredients used.

These theoretical ammonium and organic N contents were derived by using the composition of the growth medium and the α-Amino N (N_{AMINO}) and Total N (N_{TOT})contents of the ingredients as extracted from data published by the manufacturer of the culture media (Scharlau, 2011). The N_{TOT} content of the yeast extract and of the casein peptone was split into organic N (N_{ORG}) and ammonium (N_{TOT}=N_{NH4}+N_{ORG}). Simulated organic labile N includes C in its formulation, with a_C representing its C to N molar ratio. a_C has been derived from published data (Scharlau, 2011) by summation of the C to N relationship for each of the amino acids weighted according to their relative abundance (g free amino acid/g DW).

Model parameterization (Table 2) was achieved tentatively by fixing or modifying model parameters until the experimental glucose, ethanol and cell density data were reproduced reasonably well. When possible, model parameter values were taken following the work of Portell *et al*. (Portell et al., 2014). New parameters introduced here into INDISIM-*Saccha* are the respirative biomass yield (y_R), the mean individual lag time (t_{Ta}), the O_2 uptake coefficient (u_{O2}), the respirative to fermentative energy relation ($\epsilon_{R/F}$), the O_2 entrance rate to the medium (v_{O2}), the O_2 reduced by the sodium thioglycolate (f_{O2}), and the O_2 concentration at equilibrium (s_{O2}^{Eq}). The initial explored range for u_{O2} was set to 0.18-9 mmol/(g DWh) but this proved to be slightly insufficient (Table 2). A similar observation with the glucose uptake coefficient was found previously (Portell et al., 2014). $\epsilon_{R/F}$ was fixed assuming energetic production of Eqs. 1 and 2. Since, in the experimental findings, differences in the initial O_2 level were due to the molecular oxygen reduced by the sodium thioglycolate, f_{O2} was set to 60 % as experimentally observed. s_{O2}^{Eq} was set according to the O_2 solubility in water at 27 °C.

Table 1: INDISIM-*Saccha* initialization values used in the simulations.

Description	**Units**	**Value**
Time step length	h	0.1
Spatial cells per dimension	SC[c]	5
Simulated volume	ml	0.001
Initial glucose	g/l	10
Initial organic N	g/l	0.77
Initial NH_4	g/l	0.60
Initial O_2	g/l	0.0065[a]; 0.0039[b]
Organic N composition		$C_{4.7}N$
Initial sodium thioglycolate	g/l	0[a]; 0.463 [b]
Initial individuals	cells/ml	50200[a]; 30250[b]

[a] for aerobic conditions.
[b] for microaerophilic conditions.
[c] Spatial Cells.

Values from the simulated glucose depletion, ethanol formation, and population growth curves were obtained and compared against their homologous experimental points (Fig. 1). The length of both aerobic and microaerophilic experiments was the same, approximately 30 hours. Microaerophilic culture was not able to consume all the initial glucose and thus, there are no experimental (or simulated) values for low glucose levels under these experimental conditions (Fig. 1). Agreement between simulated and experimental data in two experimental situations highlights the potentiality of the approach followed to reproduce this kind of yeast batch cultures growing in varying oxygen levels. The lower maximum number of individuals reached in the simulated values is likely due to the fact that not all the initial carbon and energetic substrates can be perfectly characterized due to the nature of the media ingredients used. Hence, the yeast may be able to use many substrates as an energy and carbon source (e.g., trehalose, glycogen, glucans and manans traces) besides glucose and organic N.

4. Conclusions

The simulated results shown suggest that the approach described here has the potential to reproduce batch cultures of respiring *S. cerevisiae* cells. Hence, the contribution in question also represents a further step to obtain a microbial IBM to account for the whole set of metabolic alternatives experienced by *S. cerevisiae*. The approach would however benefit from tightly controlled experimental initial conditions and designs, following a high number of both individual and system-level determinations, and thus supporting a fruitful parameterization and calibration process. Given that the microbial IBM methodology is still being developed and the fundamental and industrial importance of *S. cerevisiae*, having well-calibrated computational models reproducing the dynamics of yeast populations under the fermentative, respirative, and respirofermentative conditions would represent a significant advancement in the field of microbial ecology.

Table 2: INDISIM-*Saccha* parameter values used in the simulations. Standard deviation of the parameters taken into account is shown between parentheses.

Parameter description	Symbol	Units	Value	Reference
Yeast C/N molar ratio	a_{yc}	mol C/mol N	6	See Portell et al. (2014)
C/N to uptake carbon sources	r_C	mol C/mol N	6	See Portell et al. (2014)
C/N to uptake nitrogen sources	r_N	mol C/mol N	12	See Portell et al. (2014)
Respirative biomass yield	y_R	g DW/g glucose	0.6	Derived from Alexander and Jeffries (1990); Thierie (2004)
Fermentative biomass yield	y_F	g DW/g glucose	0.06	Derived from Portell et al. (2014)
Potential chronological lifespan	t_d	h	1000 (0)	See Portell et al. (2014)
Maximum lifespan	h_l	scars	30	See Portell et al. (2014)
Mean individual lag time	t_a	h	8 (0.25)[a]; 6.5 (0.25)[b]	Adjusted
Glucose uptake coefficient	u_{GLU}	mmol/(g DWh)	48 (0.25)[a]; 7.5 (0.25)[b]	Adjusted
Organic N uptake coefficient	u_{CN}	mmol/(g DWh)	4.5 (0.25)[a]; 2.3 (0.25)[b]	See Portell et al. (2014)
Ammonium uptake coefficient	u_{NH4}	mmol/(g DWh)	375 (0.25)[a]; 300 (0.25)[b]	See Portell et al. (2014)
O_2 uptake coefficient	u_{O2}	mmol/(g DWh)	15 (0.25)	See text
Uptake penalty per cell scar	p_s	1/Cell scar	0.0125 (0.25)	See Portell et al. (2014)
Energy penalization due to ethanol	p_{OH}	g glucose R/(g DWh)/(g ethanol/l)	0.08 (0.25)	See Portell et al. (2014)
Ethanol tolerance	p_{uOH}	g ethanol/l	140 (0.25)	See Portell et al. (2014)
Maintenance energy	e_R	g glucose R/(g DWh)	0.01	See Portell et al. (2014)
Respirative to fermentative energy relation	$\epsilon_{R/F}$	mol fermented/mol respired	18	See text
Glucose conc. to accumulate carbon reserves	l_S	g glucose/l	0.05	See Portell et al. (2014)
Minimum mass at reproduction	m_C	pg DW	5 (0.25)	Adjusted
Min. mass increase to start bud formation	Δm_{B1}	pg DW	0.5	Adjusted
Min. mass increase for budding	Δm_{B2}	pg DW	1 (0.25)	Adjusted
Minimum time for budding	Δt_{B1}	h	0.05 (0.25)	Adjusted
O_2 entrance rate	v_{O2}	1/h	12	Adjusted
O_2 reduced by the sodium thioglycolate	f_{O2}	%	60	See text
O_2 at equilibrium	s_{O2}^{Eq}	mg/l	8.45	See text

[a] for aerobic conditions.
[b] for microaerophilic conditions.

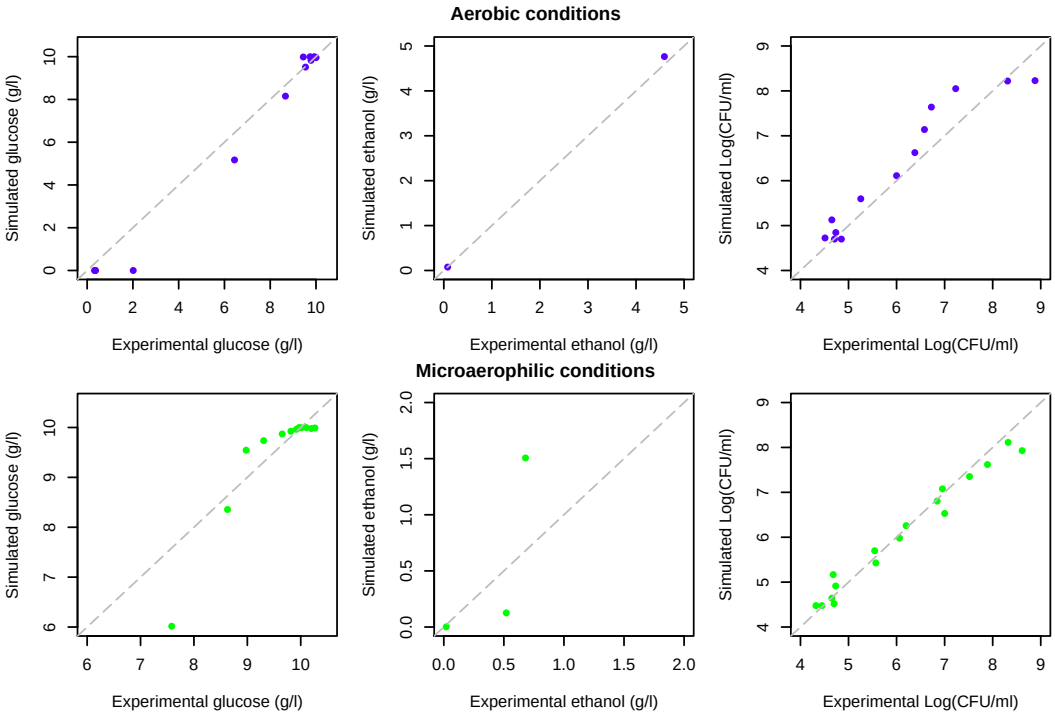


Figure 1: Experimental versus simulated glucose, ethanol, and population growth data of aerobic and microaerophilic batch cultures. Simulated values were generated using values from Tables 1 and 2. In the figure, CFU stands for colony forming units.

Acknowledgements

First author thanks Dr Clara Prats for her helpful comments about the lag phase modeling. We gratefully acknowledge the financial support of the Ministry of Education and Science (Plan Nacional I+D+i) through grant CGL2010-20160, and of the Universitat Politècnica de Catalunya.

References

Alexander, M., Jeffries, T., 1990. Respiratory efficiency and metabolite partitioning as regulatory phenomena in yeasts. Enzyme and microbial technology 12, 2–19.

Bermejo, C., Ewald, J.C., Lanquar, V., Jones, A.M., Frommer, W.B., 2011. In vivo biochemistry: quantifying ion and metabolite levels in individual cells or cultures of yeast. Biochemical Journal 438, 1–10.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S.K., Huse, G., Huth, A., Jepsen, J.U., Jrgensen, C., Mooij, W.M., Mller, B., Peer, G., Piou, C., Railsback, S.F., Robbins, A.M., Robbins, M.M., Rossmanith, E., Rger, N., Strand, E., Souissi, S., Stillman, R.A., Vab, R., Visser, U., DeAngelis, D.L., 2006. A standard protocol for describing individual-based and agent-based models. Ecological Modelling 198, 115 – 126.

Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F., 2010. The ODD protocol: A review and first update. Ecological Modelling 221, 2760 – 2768.

Gustavsson, A.K., van Niekerk, D.D., Adiels, C.B., du Preez, F.B., Goksr, M., Snoep, J.L., 2012. Sustained glycolytic oscillations in individual isolated yeast cells. FEBS Journal 279, 2837–2847.

Hellweger, F., Fredrick, N., Berges, J., 2014. Age-correlated stress resistance improves fitness of yeast: support from agent-based simulations. BMC Systems Biology 8, 18.

Portell, X., Gras, A., Ginovart, M., 2014. Indisim-*Saccha*, an individual-based model to tackle *Saccharomyces cerevisiae* fermentations. Ecological Modelling 279, 12 – 23.

Prats, C., López, D., Giró, A., Ferrer, J., Valls, J., 2006. Individual-based modelling of bacterial cultures to study the microscopic causes of the lag phase. Journal of Theoretical Biology 241, 939 – 953.

Scharlau, 2011. Handbook of microbiological culture media. Barcelona, Spain. 11 international edition edition.

Scheller, R.M., Sturtevant, B.R., Gustafson, E.J., Ward, B.C., Mladenoff, D.J., 2009. Increasing the reliability of ecological models using modern software engineering techniques. Frontiers in Ecology and the Environment 8, 253–260.

Thierie, J., 2004. Modeling threshold phenomena, metabolic pathways switches and signals in chemostat-cultivated cells: the Crabtree effect in *Saccharomyces cerevisiae*. Journal of Theoretical Biology 226, 483–501.

Urban, P.L., Schmidt, A.M., Fagerer, S.R., Amantonico, A., Ibanez, A., Jefimovs, K., Heinemann, M., Zenobi, R., 2011. Carbon-13 labelling strategy for studying the ATP metabolism in individual yeast cells by micro-arrays for mass spectrometry. Mol. BioSyst. 7, 2837–2840.

Walker, G.M., 1998. Yeast physiology and biotechnology. John Wiley & Sons.

Chapter 6

INDISIM-*Saccha* meets NetLogo

6.1 Introduction

INDISIM-*Saccha* had been implemented in the Fortran programming language. In spite of computational efficiency and other benefits, this may impose a substantial barrier for any scientist not familiar using general purpose languages to develop computer programs. This limits INDISIM-*Saccha* visibility, making global collaboration between scientists more difficult.

Utilization of the widely used, free, and open source IbM software platform NetLogo would allow INDISIM-*Saccha* to benefit from a number of advantages (see section 1.5.4) and could potentially improve the visibility and usability of the model. Nevertheless,the complexity of the model makes the implementation of INDISIM-*Saccha* in its entirity in NetLogo impractical, at least for research purposes.

Nonetheless, implementation of a streamlined version of INDISIM-*Saccha* in NetLogo would certainly take advantage of the benefits offered by this platform. At the same time, this implementation can be seen as a prototype allowing initial testing of some general programming strategies used in INDISIM-*Saccha*. For the sake of simplicity, this streamlined version has been named INDISIM-YEAST-NL, but it shares important traits with INDISIM-*Saccha*, with the use of experimental (instead of simulation) units being the most relevant. INDISIM-YEAST-NL was introduced to the scientific community at the 5^{th} International Conference on Environmental, Industrial and Applied Microbiology (BioMicroWorld 2013) and recently published in the form of the conference proceeding below (Portell *et al.*, 2014c).

6.2 The microbial individual-based model INDISIM-YEAST ready to be used in the free access NetLogo modelling environment

Publication details

Authors:	**Xavier Portell**, Anna Gras, Clara Prats, and Marta Ginovart
Title:	The microbial individual-based model INDISIM-YEAST ready to be used in the free access NetLogo modelling environment
Ref.:	In: *Industrial, medical and environmental applications of microorganisms: current status and trends*. Ed. by A. Méndez-Vilas. Wageningen Academic Publishers, p. 686. 2014

The microbial individual-based model INDISIM-YEAST ready to be used in the free access NetLogo modelling

X. Portell[1], **A. Gras**[1], **C. Prats**[2] and **M. Ginovart**[3,*]

[1] Department of Agri-Food Engineering and Biotechnology, Universitat Politècnica de Catalunya, Campus Baix Llobregrat, Esteve Terradas 8, 08860-Castelldefels, Barcelona, Spain
[2] Department of Physics and Nuclear Engineering, Universitat Politècnica de Catalunya, Campus Baix Llobregat, Esteve Terradas 8, 08860-Castelldefels, Barcelona, Spain
[3]Department of Applied Mathematics III, Universitat Politècnica de Catalunya, Campus Baix Llobregat, Esteve Terradas 8, 08860-Castelldefels, Barcelona, Spain
*Corresponding author: e-mail: marta.ginovart@upc.edu, Phone: +34 935521133

Individual-based modelling is beginning to establish itself as a viable alternative for investigating microbial systems. The individual-based model INDISIM-YEAST deals with the evolution of a yeast culture, a population with mother cells of different genealogical ages and daughter cells, by setting up rules of behaviour for each yeast according to its own rules and characteristics. It takes into account the uptake, metabolism, budding reproduction and viability of the yeast cells over a period of time in the bulk of a liquid medium, occupying a closed space with different substrates. A revised and parameterized version of INDISIM-YEAST has been implemented in the widely used, free, open source software platform NetLogo. This provides full access to this simulation model for interested microbiologists and enables them to explore the effects of inoculum characteristics on the first stages of a growing yeast population *in silico*, taking into account the great influence of the discrete and asymmetrical nature of yeast cellular division. In addition, this increases its overall understandability.

Keywords microbial individual-based model; yeast population; computer simulation; NetLogo

1. Introduction

The study with mathematical, computational, and engineering methods of microbial systems in order to understand the roles and mechanisms of the interactions of different components of these systems, and to examine the consequences of these relationships, represents a challenging, complex, scientific, and technological problem which requires collaboration among biologists, chemists, mathematicians, and others. In certain contexts, computational models offer the potential to reduce some experimental assays required when exploring new substrates and/or process options. Various researchers have postulated that simulation models are a new kind of experimental system [1].

A relatively new approach to model microbial systems, referred to as individual-based modelling (IBM), also known as biological agent-based model, is beginning to establish itself as a viable alternative for their study [2,3]. IBMs are conceptually easy to understand and do not require mastery of advanced mathematical theories. These computational models require distinct simplifications from those assumed by continuous models, and they are a good modelling alternative to deal with specific aspects related to microbial systems. In these models, the microbial cells that make up the system are treated as autonomous and discrete living entities, focussing on their characterization by means of rules of behaviour, which allow the microbes to interact, directly and/or indirectly, with one another and with the environment in which they are evolving. They support the idea that a mechanistic and scale-dependent description of microbial activity, with control of microbial biomasses, cellular cycles and their relationships with nutrients and end products is essential when dealing with the dynamics and structures of these microbial populations.

Microbial IBMs have characteristics and essential designs that cannot be described by means of equations and parameters with a synthetic formulation. They do not have a globally accepted common language for their communication, and are usually too broad to be completely described in a single informative or scientific publication [4,5]. Without the option of having a complete computational code for the simulation model, its description is mainly verbal (text) to transmit the information contained in it. Solving the problem of how to communicate IBMs may increase their scientific credibility [6]. This type of simulation model can be more difficult to analyse, understand, and communicate than traditional analytical or more classical models. Programmable modelling environments specifically designed for the implementation of IBMs can help in this communication process. In an attempt to minimize the task of implementation and execution of simulations, and at the same time to unify criteria and simplify the construction process of this type of models, universities and research centres have been developing different computing environments or platforms. Among those available today worthy of mention are Swarm, Echo, Repast, MASON, and NetLogo. Although from among those available the choice of an agent-based modelling toolkit is not straightforward [7], NetLogo [8] offers attractive

possibilities. Netlogo is a computing framework which enables the simulation of natural and social phenomena and is continually being updated by the Center for Connected Learning and Computer-Based Modeling in the United States. It can coordinate all the instructions given to thousands of agents or individuals so that they all operate independently among themselves and with the environment. It is a free tool accessible on the Web, and its web page contains extensive documentation, manuals and tutorials.

The microbial IBM known as INDISIM-YEAST deals with the evolution of a yeast culture, a population with mother (or parent) cells of different genealogical ages and daughter (or virgin) cells, by setting up rules of behaviour for each yeast according to its own characteristics, and taking into account the uptake, metabolism, budding reproduction, and viability of these cells, over a period of time in the bulk of a liquid medium and occupying a closed space with different kinds of substrates [9-11]. INDISIM-YEAST previous versions use the Fortran programming language but sometimes this can be a barrier for non-computer scientists. Usually, verbal model descriptions are provided for this computational model and it is difficult to specify unambiguously all the details of the different submodels implemented in the computer code. Thus, it is good practice to provide executable versions that can accomplish the simulation results [12] or, better still, to make the computer codes accessible in a suitable environment for reproducibility and analysis of the simulation outputs.

The aim of this work is to present a revised and parameterized version of the INDISIM-YEAST model implemented in the widely used, free, open source software platform NetLogo, providing full access to this simulation model for interested microbiologists, which allows them to explore the effects of inoculum characteristics on the first stages of a growing yeast population *in silico*, taking into account the great influence of the discrete and asymmetrical nature of yeast cellular division. In addition, this increases the overall understandability of this model and ensures complete information transfer.

2. The model INDISIM-YEAST

In recent years, many researchers have attempted to model yeast fermentation and different approaches have been considered. However, even segregated and structured models do not consider the complete cell cycle dependence or physiological properties of different cells as model variables. A direct consequence is that these kinds of models are only valuable to describe and predict cultures with homogeneous and constant population composition rather than transient behaviour or heterogeneous compositions.

To our knowledge INDISIM-YEAST is the first work that models the processes involved in yeast population growth using IBM simulations. A first general presentation of the main ideas of this microbial IBM, INDISIM-YEAST, was published in 2007 [9]. This model is a descendant of the INDISIM model family [13]. It is assumed that a yeast population grows in the bulk of a liquid medium where we consider variables that are space and time dependent. These variables control the amount of abiotic components, and are identified as glucose, the nutrient particles, and ethanol and CO_2, the end product particles arising from the yeast metabolism and excreted to the environment. The time evolution of the population is divided into equal intervals associated with computer steps or time steps. The spatial domain where the system evolves is a grid made up of spatial cells. In batch culture the medium is not altered by further nutrient addition or removal. At each time step, a single organism, an individual yeast cell, is defined by a set of time-dependent variables, which describe and control its individual properties. For each microorganism that forms the population INDISIM-YEAST implements a set of rules for the following actions: motion, uptake and metabolism of nutrient, budding reproduction and viability. We identified its position in the spatial domain; its biomass; its genealogical age as a number of bud scars on the cellular membrane; the reproduction phase of the cellular cycle, unbudded or budding phase; its “start mass”, the mass to change from unbudded to budding phase; the minimum growth of its biomass for the budding phase; the minimum time required to complete the budding phase; its survival time without satisfying its metabolic requirements. Random variables are used to characterize the yeast cells and their individual actions.

In 2008,with the aim of furthering the skills associated with the use of this individual-based simulator, a website from which INDISIM-YEAST can be executed was developed and presented [12], but without allowing full access to the computer code. Although the application resulted in a very versatile program that could be used in controlled simulation experiments via Internet, it permitted only a limited interaction with this model.

Practicable possibilities of this yeast IBM with some simulated results have been presented recently. For instance, taking into account the yeast ageing and inoculum size are factors that affect industrial fermentation (particularly with reused yeast in the production of beer, “serial repitching”), INDISIM-YEAST was used to explore the effects of inoculum size and ageing on the first stages of the dynamics of yeast population growth [10]. Only an IBMs such as this allows the study of small, well-characterized, microbial inocula. Several simulations were performed to analyse the effect of inoculum size and genealogical age of the yeast cells forming the inoculum on the lag phase, first division time and specific growth rate. To complete this study, the effects of these two factors on the dynamics of virtual yeast fermentations were also explored, focusing mainly on: i) the first stages of population growth, ii) the mean biomass evolution of the population, iii) the rate of

glucose uptake and ethanol production, and iv) the biomass and genealogical age distributions [11]. The ultimate goal was to integrate these results in order to make progress in the understanding of the composition of yeast populations and their temporal evolution in beer or other fermentations. Simulation and experimental results showed that there is a clear influence of these initial features of the inocula on the subsequent growth dynamics.

Although for a complete description of different parts of INDISIM-YEAST the reader can see some previously published papers, a formal and general description of this simulation model was presented adopting the ODD (Overview, Design concepts, Details) standard protocol in the work of Ginovart and co-authors [11]. This ODD protocol is widely accepted for describing IBMs and contributes to the unification of their formulation and implementation [4,5].

Nevertheless, we firmly believe that a really useful way to analyse this yeast simulator is to experiment and explore the manner in which it reacts to changes in parameter values, initial conditions or assumptions made in its design, which is only possible if we can have the computer code and the possibility to manipulate it. Without the complete code only limited handling and analysis of this computational model can be performed.

3. The simulator INDISIM-YEAST in NetLogo

To construct INDISIM-YEAST, a world is created in which the programmer (or the future user) has access to all of the laws and elements of that world, and can also manipulate all the relationships among its abiotic and biotic components. Equipping INDISIM-YEAST with graphical user interfaces makes the simulator friendlier and easier to understand and so test the model, and this is possible with its implementation in the NetLogo modelling environment. Figures 1-3 show the three typical tabs that NetLogo platform provides. Fig. 1 shows the interface tab of this preliminary version of INDISIM-YEAST simulator, and Figs. 2 and 3 show the information tab and the code tab respectively.

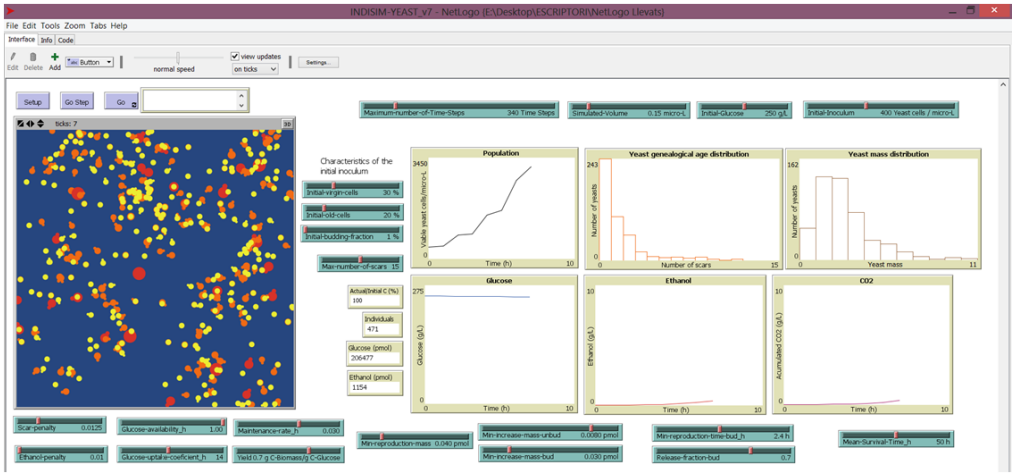


Fig. 1 INDISIM-YEAST user interface. The sliders allow the changing of initial values to the simulation, model parameters and characteristics of the initial inoculum. There are buttons to setup and run the simulation at the top. On left there is the lattice representation of the modelled system.

This simulator includes individuals (named *turtles* in NetLogo) which are yeast cells and, in order to visualize changes within the structure of the population throughout the time, they are represented according to their number of scars following a colour code (Fig. 1). The three yeast groups are: i) daughter or virgin cells with 0 scars (yellow), ii) middle-aged cells with 1 <= #Scars <= 5 (orange), and iii) old cells with #Scars > 5 (red). It is also possible to distinguish between the unbudded and budded cells, symbolised by dots and dots with a protuberance, respectively (Fig. 1). A two dimensional array or lattice of 30x30 grid cells (named *patches* in NetLogo) represents the simulated space in which the culture medium is contained. The volume, in µL, can be fixed in the interface window by the user. Observations are provided with monitors and plots, displaying the yeast concentration (number of the viable yeast cells·μL^{-1}), the genealogical age distribution, the microbial biomass distribution, and amounts of the substrates considered, glucose (gL^{-1}), ethanol (gL^{-1}), and CO_2 (gL^{-1}), over time (Fig. 1). The simulations provide population (global), properties, as well as properties that pertain to individual yeast cells (microscopic properties). NetLogo user interface makes it easy to change initial values using the sliders (Fig. 1), such as the percentages of the virgin (or daughters) cells, middle-aged cells and old

cells that make up the initial inoculum to start the growth of the population, the initial amount of glucose, or other model parameters [9-11].

For the first parameterization of this simulation model the study carried out by Portell and co-authors [14] is being used as the main reference and basis for this project. Since the predominant data available comes from experimental assays designed to achieve the mean representative value to characterize the population behaviour, assigning values to parameters in this individual-based framework is not a straightforward issue. Values obtained in this way are likely to include interactions between individuals as well as between individuals and the specific experimental system used. IBM output, which is already the result of the parameter values and stated interactions, is intrinsically and inherently affected by this biased data [14].

Inputted parameters in the simulator must be linked with the simulation units. To make practicable the use of this simulator the input data should be expressed in those units more frequently used by experimentalists. Nevertheless, to deal with this IBM and to establish the stoichiometric reactions for the individual metabolism it is more suitable to consider molar units. Thus, the composition of the culture medium, the biomasses and other pertinent model parameters are expressed in picomols in the computer code, so an initial unit conversion is performed. The processes are modelled discretely with time or computer steps that represent five minutes approximately. The total duration of the simulation can also be selected by the user.

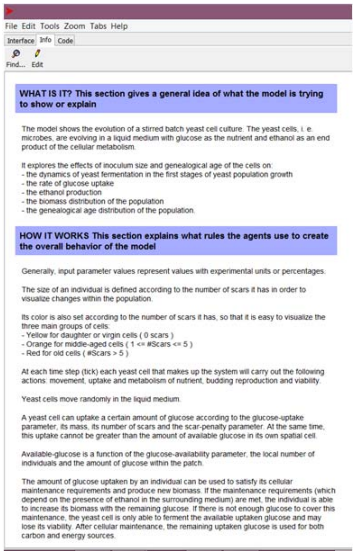


Fig. 2 A screenshot of the INDISIM-YEAST information tab containing some ideas regarding the simulation model and its implementation. The parts that can be found correspond to: What is it?, How it works, How to use it, Things to notice, Things to try, Extending the model, NetLogo Features, Related models, and Credits and references.

4. Final remarks

The main motivation to implement INDISIM-YEAST in NetLogo programming environment was not to develop a predictive simulator or adjust some experimental data, but to provide a preliminary simplified and parameterized version of this model to explore the dynamic and structure of yeast populations growing in batch culture with different types of initial inocula or glucose concentration. By contrasting both the individual (yeast cells) and global (population or simulated culture) properties, it is expected that we can gain insight into the interrelation between these two types of data, which we hope will help to deal with the macroscopic behaviour observed in experimental yeast research. This simulator can facilitate understanding of the underlying microbial concepts, structure and algorithms of the INDISIM-YEAST model, and of microbial IBMs in general.

Given the NetLogo's fairly flat learning curve and very complete documentation about this modelling environment, even users without extensive modelling and/or computing experience can modify the computer code and, therefore, investigate or implement alternative mechanisms for additional processes presumed relevant to a particular microbial study approach. The simulator, which is still in the improvement phase, can be requested directly from the authors.

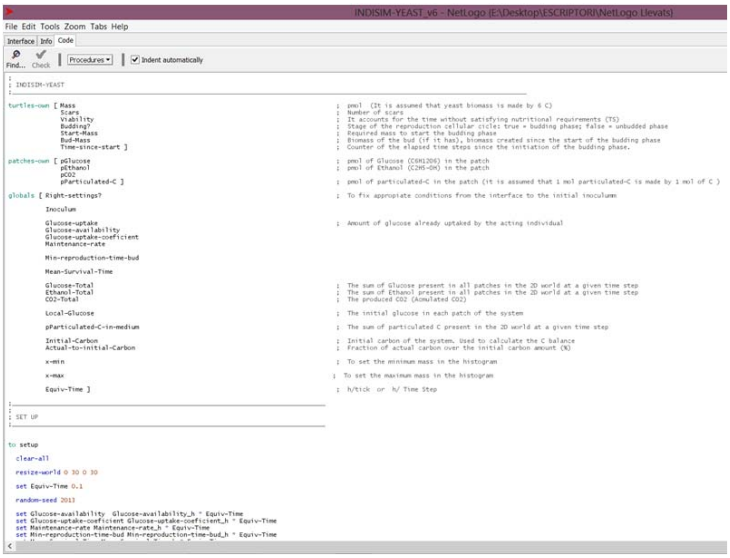


Fig. 3 A screenshot of the INDISIM-YEAST procedures tab where the computer code can be found.

Acknowledgements Financial support of the Spanish Government (MICINN, CGL2010-20160) is gratefully acknowledged.

References

[1] Peck SL.Simulation as experiment: a philosophical reassessment for biological modeling. *TRENDS in Ecology and Evolution*. 2004; 19:530-534.

[2] Grimm V, Railsback SF. *Individual-based modelling and ecology*. Princeton and Oxford: Princeton University Press; 2005.

[3] Hellweger FL, Bucci V. A bunchof tiny individuals - individual based modeling for microbes.*Ecological Modelling*. 2009; 220:8-22.

[4] Grimm V, Berger U, Bastiansen F, Eliassen S, Ginot V, Giske J, et al. A standard protocol for describing individual-based and agent-based models. *Ecological modelling*. 2006; 198:115-126.

[5] Grimm V, Berger U, DeAngelis, DL, Polhill JG, Giske J, Railsback SF. The ODD protocol: A review and first update. *Ecological Modelling*. 2010; 221:2760-2768.

[6] Railsback SF, Grimm V. *Agent-Based and Individual-Based Modeling: A Practical Introduction*. Princeton and Oxford: Princeton University Press; 2011.

[7] Devillers J, Devillers H, Decourtye A, Aupinel P. Internet resources for agent-based modelling. *SAR and QSAR in Environmental Research*. 2010, 21:337-350.

[8] Wilensky U, 1999. NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL. Available at http://ccl.northwestern.edu/netlogo/ Accessed October 11, 2013.

[9] Ginovart M, Xifré J, López D, Silbert M. INDISIM-YEAST, an individual-based model to study yeast population in batch cultures. In: Méndez-Vilas A, ed. *Microbiology Book Series nº 1 Vol. 1, Communicating Current Research and Educational Topics and Trends in Applied Microbiology*, Badajoz: Formatex; 2007: 401-409.

[10] Ginovart M, Prats C,Portell X, Silbert M. Exploring the lag phase and growth initiation of a yeast culture by means of an Individual-based Model. *Food Microbiology*. 2011; 28:810-817.

[11] Ginovart M, Prats C, Portell X, Silbert M. Analysis of the effect of inoculum characteristics on the first stages of a growing yeast population in beer fermentations by means of an Individual-based Model. *Journal of Industrial Microbiology and Biotechnology*. 2011; 38:153-165.

[12] Ginovart M, Cañadas JC. INDISIM-YEAST: an individual-based simulator on a website for experimenting and investigating diverse dynamics of yeast populations in liquid media. *Journal of Industrial Microbiology and Biotechnology*. 2008; 35:1359-1366.

[13] Ginovart M, López D, Valls J. INDISIM, an individual based discrete simulation model to study bacterial cultures. *Journal of Theoretical Biology*. 2002; 214:305-319.

[14] Portell X, Gras A, Ginovart M. INDISIM-*Saccha*, a parameterized Individual-based Model to tackle *Saccharomyces cerevisiae* batch cultures. *Submitted to Ecological Modelling*. 2013.

Chapter 7

General discussion

INDISIM-YEAST has proven its ability to distinguish differences in the evolution of the population that emerges from small inocula with different age structures (chapter 3). Although no published experiments perfectly matching the virtual simulation series conducted have been found in the literature, the patterns and tendencies of the simulated outputs are in broad agreement with experimental research related to *S. cerevisiae* under fermentative conditions. It seems clear that the model is able to capture relevant mechanisms of the modelled system, at least for the outputs assessed. Hence, the simulator INDISIM-YEAST can be regarded as an appropriate starting point for development of further *S. cerevisiae* IbMs.

Budding reproduction leading to an asymmetric cell division is a determining factor in the genealogical age structure of the population which, in turn, is closely related to the size of the cells due to the ageing process. Consequently, individual-size distributions are advisable inputs of the parameterization and calibration of a yeast IbM. Chapter 4 reports on, and describes, the differences in the central tendency, variability and distribution shape of the cell diameter of *S. cerevisiae* populations growing in batch cultures. Different high throughput experimental techniques (electric particle analysis, light diffraction, and flow cytometry) have confirmed stated differences and complemented each other. Aerobic and microaerophilic grown cultures have shown differences in cell diameter (electric particle analysis, and light diffraction), cell size (forward scatter flow cytometry), and cell granularity (side scatter flow cytometry).

Carrying out the analysis and comparison of different experimental techniques gives the modeller valuable experimental knowledge and background to design μIbM-oriented experiments. Thence, future interdisciplinary work, a must in the advocated μIbE (Kreft *et al.*, 2013), is indirectly benefited.

INDISIM-YEAST outputs tested in chapter 3 can be regarded as being mainly influenced by the reproduction submodel and the cell ageing process. Given the simplicity of the metabolism modelled in the INDISIM-YEAST, and the positive qualitative agreement found between simulated and experimental published results, it would seem rational to develop an IbM of *S. cerevisiae* focusing on yeast metabolism, and specifically the sugar catabolic alternatives

which are characteristic of this yeast. In addition to the yeast metabolism to be introduced, INDISIM-YEAST significantly surpassing these capabilities requires a fully-prepared model to be compared against experimental data. The simulator developed has been named INDISIM-*Saccha* and was introduced in chapter 5.

First application of INDISIM-*Saccha* analysed the dynamics of *S. cerevisiae* fermentations, which means that an aerobic metabolism is assumed (section 5.2). This first application introduced, parameterized, calibrated, evaluated the model adequacy, and used the model to assess *in silico* ethanol production by means of virtual experiments. The calibration procedure, and the performance and analysis of the data from the virtual experiments has been undertaken using the statistical programming language R. This involved the development of scripts fitting the functionalities required and also the use of existing R packages. Parallel computing techniques were used in the script developed to automate the realization of simulations. Model description was conducted following the most accepted protocol to describe IbMs, the ODD protocol. Parameterization of the model was successfully achieved considering the positive assessment of the model adequacy conducted testing several model predictions both at a system level (glucose depletion, population growth curves) and single-cell level (fraction of budded cells, genealogical age distribution, and cell diameter distribution evolutions). Individual cell size data from chapter 4 (section 4.2) played an important role in this assessment. Parameterization of this μIbM was not so direct and straightforward a process as that of the translation from available published data to individually-based parameters of the model, demanding a thorough knowledge of yeast biology, experimental practices, and modelling expertise. The parameterization and the categorical calibration of the model allow population growth, glucose depletion (closely linked to ethanol production), and temporal evolutions of ages and sizes (masses) of the yeast cells making up the population to be simulated satisfactorily and well enough to be used in the intended prospective study. The iterative calibration process performed was capable of reducing uncertainty from the parameterization process on some of the model parameters. Although this reduction was sufficient to tackle this first stage of study, further utilization of the model calls for the use of a greater variety of both population and single-cell observations, which would also greatly improve our confidence in the assumptions made throughout the modelling process. Results of the virtual experiments conducted suggest that differences in the structure of the initial population, and specifically in the cell size distribution, can drastically affect the performance and productivity of fermentations, and encourage routine characterization of the inocula in the biotechnological industry. The implemented model is capable of dealing with this structural variability, thus encouraging us in further development and more practical applications of INDISIM-*Saccha*.

INDISIM-*Saccha* has also been adapted to take into account the aerobic growth of *S. cerevisiae* and contrasted with two experimental trials with different oxygen levels in the medium. The preliminary simulated results achieved with the model suggest that the approach holds potential to reproduce aerobic batch cultures of *S. cerevisiae*. Hence, this also represents a further step in obtaining a μIbMwhich will account for the whole set of metabolic alternatives

experienced by *S. cerevisiae*: fermentative, respirative, and respirofermentative.

The main advantage of a simulator implemented in the procedural programming language Fortran is the computational efficiency achieved, clearly a point to be taken into account for a μIbM. Conversely, however, the considerable specific knowledge required to manipulate the INDISIM-*Saccha* simulator developed makes its utilization difficult for non-experienced external users. The visibility and usability of the methodology would then benefit from the utilization of a freely available and widely used high-level agent-based modelling environment such as NetLogo. Bearing this in mind, a streamlined version of INDISIM-*Saccha* has been implemented and presented in chapter 6. The simulator obtained, named INDISIM-YEAST-NL for practical reasons (see section 6.1), provides a simplified and parameterized version of the model allowing the dynamic and structure of yeast populations growing in batch culture with different types of initial inocula or glucose concentration to be explored. Given NetLogo's fairly flat learning curve and very complete documentation about this modelling environment, even users without extensive modelling and/or computing experience can modify the computer code and, therefore, investigate or implement alternative mechanisms for additional processes presumed relevant to a particular microbial study approach. From INDISIM's perspective, INDISIM-YEAST-NL can facilitate understanding of the underlying microbial concepts, structure, and algorithms of INDISIM-*Saccha*, and of μIbM in general. From the didactic point of view, this implementation clearly facilitates understanding of the relationship between individual (yeast cells) and global (population) properties; and may enhance comprehension of underlying microbial concepts such as batch culture, inoculum size, yeast ageing, and alcoholic fermentative processes, and thus reinforces and consolidates the μIbM methodology.

Chapter 8

Conclusions

1. A number of simulators: INDISIM-YEAST, INDISIM-*Saccha*, and INDISIM-YEAST-NL which allow the study of microbial systems driven by the yeast *S. cerevisiae* have been designed, implemented and already used in the current dissertation.

2. The adequacy of the INDISIM-YEAST model as a core model for the IbM INDISIM-*Saccha* development has been assessed by realization of controlled simulation series and comparison of the results against the current knowledge available in specialized experimental literature. The analysis performed has led to a positive assessment of the adequacy of the model.

3. The INDISIM-YEAST analysis conducted mimicked a brewery practice called serial repitching, and investigated: the lag and exponential phases, the effects of the inoculum size and the genealogical age on interesting system-level and individual-based outputs. The positive assessment of the model, and the high influence of the reproduction submodel and ageing process on the evaluated outputs has highlighted the validity of these model parts, and highlighted a need to focus on further developing the metabolism submodel.

4. Differences between growth phases in the central tendency, variability, and distribution shape of the cell diameter of *S. cerevisiae* population growing in batch cultures have been reported. In short, diameter distribution (assessed by electrical particle analysis) shifts to the right until the exponential phase is reached, and then shifts to the left again, which results in a lower mean cell size once in the stationary phase. The distribution shape evolves from an initial triangular distribution to become slightly positively skewed, something not found in the stationary phase, where a more symmetrical distribution is recorded. Flow cytometry data has revealed additional heterogeneity in the cells once the stationary phase has been reached, allowing two subpopulations of cells to be distinguished.

5. Differences in the central tendency, variability, and distribution shape of the cell diameter of *S. cerevisiae* stationary phase populations growing in two culture conditions has also been reported. Cell diameter distribution (assessed by electrical particle analysis) under microaerophilic conditions is situated to the right of its equivalents under aerobic conditions, subsequently showing a greater mean cell size. Nevertheless, progression thorough the stationary phase reduces this difference. Flow cytometry has also revealed changes between growth conditions in the heterogeneity of the cells making up the population, with subpopulations of cells clearly distinguished in aerobic conditions but highly homogeneous in microaerophilic conditions.

6. Utilization of a complexity index in order to summarize and obtain information about individual-based distributions has, for the first time, been used jointly with experimental techniques for electrical particle analysis, light diffraction, and flow cytometry. The use of this heterogeneity index has confirmed the heterogeneity of these yeast populations thoroughout its development, and may help to differentiate individual-based distributions from different growth phases of the culture. This also suggests that a complexity index may be useful in comparing the evolution of simulated and experimental individual-based distributions.

7. A quantitative IbM to study the dynamics of *S. cerevisiae* fermentations has been designed, implemented in Fortran 90 language, and termed INDISIM-*Saccha*. Description of the model developed has been conducted following the most accepted protocol to describe IbMs, the ODD protocol, which favours reproducibility of the simulation work, and therefore lends credibility to the model.

8. The developed model once parametrized mainly using available external published data, has been categorically calibrated using glucose and total cell numbers data from two experimental conditions by means of an iterative procedure developed and conducted in the free software environment R.

9. The adequacy of the model after the calibration step has been assessed by testing several model predictions both at a population and individual-level. Individual cell size data from the electric particle analysis shown previously in this thesis has played an important role in this assessment.

10. The IbM developed has been used to conduct virtual experiments whose results suggest that differences in the cell size distribution of the initial population can drastically affect the performance and productivity of fermentations, and encourage routine characterization of the inocula in the biotechnological industry.

11. INDISIM-*Saccha* has been extended to analyse the dynamics of *S. cerevisiae* cultures evolving in a liquid medium with oxygen by introducing the two sugar catabolic alternatives characteristic of this yeast: fermentative and respirative catabolism. Resulting

methodology represents a further step to obtain a μIbM accounting for the whole set of metabolic alternatives experienced by *S. cerevisiae* cells: fermentative, respirative, and respirofermentative metabolism.

12. Comparison of glucose, ethanol, and cell density evolutions collected in two experimental conditions differing in their oxygen level against their homologous simulated points, arising from the extended model, suggests that the approach holds potential for reproducing and investigating aerobic batch cultures of *S. cerevisiae*.

13. The ability INDISIM-*Saccha* to link population structure to macroscopic observations will clearly gain practical importance in the near future due to the likely availability of high-throughput devices, affordable for high and low added value bioindustries alike.

14. Realization of the controlled simulation series and virtual experiments of this contribution has required development of tools implemented in the software R. These tools can be easily modified in order to undertake any inverse modelling technique, analysis step (e.g., a sensitivity or uncertainty analysis), or virtual experiments should these be further required. Said tools rely on multithreading technology to automatize simulations, and thus allow advantage of the current computing architectures to be taken, and computing time to be saved. This is especially relevant in the field of the μIbMs where a high number of individuals have to be routinely simulated.

15. The implemented INDISIM-*Saccha* model(s), along with the developed R tools can be seen as an effective and promising IbM framework to quantitatively study *S. cerevisiae* systems.

16. INDISIM-YEAST-NL, a streamlined version of INDISIM-*Saccha*, has been designed and implemented in the freely available programming environment NetLogo. The platform documentation and code availability increases the comprehensibility and completeness of the model description, and makes the model easily accessible even to non-experienced users, for didactic purposes as well as for modifications and extensions. Therefore, INDISIM-YEAST-NL lays the foundations for a deeper understanding of the developed methodology and μIbMs in general, and will facilitate future interactions with potential INDISIM-*Saccha* external users.

Chapter 9

Perspectives and future work

Further work with the modelling approach developed would benefit from tightly controlled initial experimental conditions and designs, and from following a high number of both individual and system-level determinations. This valuable data would support fruitful parameterization and calibration processes. Given that μIbM methodology is in its infancy and the fundamental and industrial importance of *S. cerevisiae*, having well-calibrated computational models — that reproduce the dynamics of the yeast across the whole range of experimental conditions to which the yeast was exposed, and which may cause a fermentative, respirative or respirofermentative individual metabolism to be originated — would represent a significant advancement in the field of microbial ecology, and eventually for industrial microbiology. Achieving this would likely involve a closer relation between experimentation and individual-based modelling, which is seen as a promising opportunity to gain a better understanding of microbial systems (Kreft *et al.*, 2013).

Calibration and a genuine increase in the external confidence of a complex IbM ideally requires following microbial cultures with a high number of both individual and population-based determinations. For INDISIM-*Saccha*, in the IbM developed in the present thesis, in addition to glucose, ethanol and population growth curves, the O_2, NH_4, and organic N medium concentrations, the inclusion is proposed of cell size (including, at least, the inoculum, and exponential, and stationary phase distributions) and scar distributions as well as the fraction of budding cells throughout the duration of the experiment. Additional determination of the experimental mean biomass yield would allow a finer parametrization of the maintenance related model parameters (p_{OH} and e_R[1]) and the biomass yields used (y_R and y_F). The entrance rate of O_2 into the system could first be experimentally measured ensuring anaerobic conditions (e.g. using pure CO_2 in the inlet gas) and then the O_2 evolution measured once the inlet gas composition changed. In experimental conditions however, it is likely that the CO_2 produced by the yeast will greatly reduce the measured value. This can be overcome by modifying the experimental set-up to ensure removal of the CO_2 produced by the yeasts and a constant O_2

[1]See appendix C for parameters description

concentration in the gaseous phase. The methodology would also benefit from using defined media, allowing firm control of the initial experimental conditions, and isolation or delimiting of the metabolic events taking place within the cell. All this would undoubtedly result in a simulator perfectly prepared to achieve interesting, potentially unprecedented, knowledge of the complex *S. cerevisiae* microbial systems, and adequate for use in more direct and applied biotechnological contexts.

Musalais is a naturally fermented alcoholic beverage made with grapes that has been prepared by the Uygur minority people in the Xinjiang province of China for more than 2,000 years. Musalais is produced by a traditional process differing from other conventional methods of wine production (Lixia *et al.*, 2012). In a nutshell, the grape juice is boiled at about 92 °C for more than fifteen hours, cooled overnight to room temperature, transferred into earthware jugs, sealed and allowed to ferment for 45 days. Cooled fermentation juice is naturally inoculated with spontaneously occurring yeasts, mainly *S. cerevisiae* strains, at a level of under one hundred individuals per ml [2]. Contributions presented in this dissertation have recently enabled contact with Dr Lixia; lecturer in the Food Science Department and member of the Research on Traditional Musalais and Yeast Population in the Musalais Spontaneous Fermentation research group at Tarim University, China; Dr Lixia is interested in using the INDISIM-*Saccha* framework to study the biological factors affecting the complex spontaneous fermentations of Musalais.

At the time of writing this document, the project entitled "Numerical Simulation of Traditional Musalais Process with Computational Fluid Dynamics and INDISIM to research its Complex Dynamic and Global Modelling of the whole process with Dynamic Bayesian Networks" has been submited by Dr Lixia's research group in response to the 2014 Chinese Government's call for research grant proposals on "National Nature Science Foundation of China". The research group where this thesis has been complete, the Discrete Modelling and Simulation of Biological Systems research group (MOSIMBIO) of the Universitat Politècnica de Catalunya (UPC), will hopefully provide the modelling support for the stated project should the Chinese Government award them a grant for the purpose. Given the experimental expertise and capabilities of Dr Lixia's group along with the INDISIM-*Saccha* framework and and the modelling expertise of the MOSIMBIO group, this project will hopefully represent an invaluable step forward in the field of microbial ecology.

[2] Dr Lixia, personal communication.

Appendices

Appendix A

INDISIM-*Saccha* description

The model description follows the ODD (Overview, Design concepts, and Details) protocol for describing individual- and agent-based models (Grimm et al., 2006; 2010). To better follow this description, the reader should bear in mind the following: (i) individual and spatial cell state variables, which record the specific state of the system at each time step, are designated with a capital letter followed by a suffix; (ii) the model parameters are designated with a lower case letter; (iii) any particular random realizations following a given statistical distribution are designated with a lower case letter with a dieresis; and (iv) the modelled organic molecule are designated as CN with a suffix. Some of the important suffixes are: GLU (glucose), OH (ethanol), NH4 (ammonium), CO2 (carbon dioxide), and MIC (from a microorganism).

A.1 Purpose

The model has been developed to analyse the dynamics of *Saccharomyces cerevisiae* anaerobic batch cultures evolving into a non-stirred liquid medium with glucose as a main C source and organic and inorganic N sources.

A.2 Entities, state variables, and scales

The model takes into account three entities: *S. cerevisiae* individual cells, spatial cells, and the environment. A yeast cell (I_i) is defined by the variables: $X(t)$, $Y(t)$, and $Z(t)$, identifying its position in the domain; $M(t)$, its structural mass (CN_{MIC}-pmol); $B(t)$, its genealogical age (bud scars); $P(t)$, the reproduction phase in the cellular cycle in which the cell is currently (unbudded or budding phase); $M_{Start}(t)$, its "Start mass" (CN_{MIC}-pmol), the mass required to change from the unbudded to the budding phase; $M_{inc}(t)$, the increased mass (CN_{MIC}-pmol) since the cell entered to the budding phase; $T_{inc}(t)$, time spent into the current reproduction phase (time steps); $R_{GLU}(t)$, the amount of C stored in the cell as reserve carbohydrates or in the model as a glucose polymers (glucose-pmol), $R_{CN}(t)$, the amount of organic N stored in the cell as a reserve (CN-pmol); $C_{GLU}^{in}(t)$, amount of non-metabolized glucose inside the cell (glucose-pmol); and, $D(t)$ the mortality index to evaluate cell viability. Letting $I=I(t)$ denote the number of individuals at time t, and identifying an individual by i, the population's state at t is:

$$I_n = \left\{ I_i \left[X, Y, Z, M, B, P, M_{Start}, M_{inc}, \mathrm{T}_{inc}, R_{GLU}, R_{CN}, C_{GLU}^{in}, D \right] \right\}_{i=1,2,\ldots,n} \quad \text{(A.1)}$$

The simulated space is a cube which is divided into spatial cells (SC) of equal size and shape (S_{xyz}) described by a vector defined by the variables: $S_{GLU}(t)$ the amount of glucose (pmol); $S_{CN}(t)$, the amount of organic N (pmol); $S_{NH4}(t)$, the amount of ammonium (pmol); $S_{OH}(t)$, the amount of ethanol (pmol); $S_{CO2}(t)$, the amount of carbon dioxide (pmol). The whole three dimensional grid is then described by:

$$Grid_{3D} = \left\{ S_{xyz} [\mathrm{S_{GLU}}, \mathrm{S_{CN}}, \mathrm{S_{NH4}}, \mathrm{S_{OH}}, \mathrm{S_{CO2}}] \right\}_{x=1,\ldots,iq; y=1,\ldots,iq; z=1,\ldots,iq} \quad \text{(A.2)}$$

The environment simulates a liquid medium closed in a cube whose faces do not allow the ingress or egress of either organic or inorganic elements. However, to account for the required periodic boundary conditions, the spatial domain is wrapped, so organic and inorganic elements moving out of the spatial domain reappear on the opposite face of the relevant axis. Substrates within the medium are subject to diffusion phenomena.

The temporal evolution of the system is divided into equal intervals associated with time steps (TS).

A.3 Process overview and scheduling

Global simulation scheduling is mainly made up of four sections: (i) the initialization of the simulated system, when the input data is entered, (ii) the setting of the initial configuration of the population, (iii) the initial setting of the space, and (iv) the time step loop which is repeated until the end of the defined time steps. The time step loop scheduling includes, chronologically: (iv.i) the random order of the individuals' acting order, which will act one behind the other, (iv.ii) the individual actions loop (sections A.7.1-A.7.5), (iv.iii) the actions over the medium (section A.7.6), and (iv.iv) output of the desired aggregated and state variables.

At each time step and at the individual actions loop, existing yeast cells perform, sequentially, the following set of actions: motion (section A.7.1), uptake (section A.7.2), metabolism (section A.7.3), reproduction (section A.7.4), and lifespan (section A.7.5). Yeast activity is locally described; therefore actions of individuals take place and only affect the spatial cell they are in. Variables are updated as soon as their value is calculated by a process (asynchronous updating).

A.4 Design concepts

Emergence. The main phenomena to emerge from the model are the evolution of the nutrient concentrations over time, the population growth behaviour, and the distribution of individual properties at particular times through the system evolution, emphasizing mass and scar distributions as a direct simulation homologous to well-established experimental techniques. The phenomena described above are not imposed by model rules but are consequences of individual behaviour and interactions between biotic and abiotic elements.

Adaptation. The model assumes that the individual yeast changes its metabolism as an adaptation strategy to the glucose depletion. The individual begins to store carbohydrates as C reserve that will be required to survive during starvation conditions (Walker, 1998).

Sensing. Individual yeasts are assumed to sense medium glucose concentration and change their internal metabolic fluxes to adapt to the

changing environmental conditions. This sensing mechanism is not explicitly modelled but individuals are simply assumed to know external glucose.

Interaction. Both intraspecific direct and indirect interaction phenomena are taking place within the modelled system. Indirect interactions include competition for nutrient sources (glucose, organic N and ammonium). Detrimental effects into individual fitness due to external ethanol content, which is a main product of the individual yeast metabolism, can be seen as a direct interaction among yeast individuals.

Stochasticity. Stochasticity is introduced into the model when setting some parameter values of the individuals using a positive truncated normal distribution. Randomness is also considered when the rules are applied to individuals and to spatial cells by using probabilistic distributions to deal with or manage individual events. Random processes or events include: (i) movement of individuals; (ii) maximum uptake rate at a given time interval; (iii) computation of the toxic effect of the ethanol on the glucose uptake; (iv) computation of maintenance energy requirements; (v) yeast start mass assignation; (vi) checking successful completion of budded phase (both within expended time and accumulated cell mass); and (vii) checking for cell death. The sequence of action of individuals changes randomly at each time step in order to avoid privileging first-acting cells.

Observation. The behaviour of the system is followed by means of (i) data related to the global properties of the system and (ii) data concerning properties of individual yeast cells. The former may include, but are not limited to, information about the temporal evolution of the amount of nutrients (glucose, organic N, and ammonium), the amount of metabolites (ethanol and carbon dioxide), the average nutrient consumption, the number of viable yeast cells, the number of dead yeast cells, viable yeast biomass, dead yeast biomass, heat dissipation of the system, maintenance energy expended by the yeast population (defined as the number of metabolized nutrient particles not used in the production of new biomass), and the mean mass of the cell population. In addition, it is feasible to obtain the evolution of microscopic population parameters, which are controlled at individual level, such as, among other potential individual properties, the distributions of genealogical age and of the mass of the individuals. These distributions are related mainly to the cellular cycles of budding reproduction, and reflect the state of the yeast population at given times in the fermentation process. Because the simulator saves information about every cell at each time step, this recorded information makes it possible to construct bar charts and histograms to represent the latter distributions. These simulation outputs show the time evolution of the structure of the population throughout the yeast culture. The preceding separation of output simulation results mirrors the classification of experimental techniques used to study said properties.

A.5 Initialization

Following the same methodology as Gras and co-workers (Ginovart et al., 2005; Gras et al., 2011), individual yeast mass is expressed in molar units, so that each mass pmol is composed by a_{yC} pmols of C and 1 pmol of N. Additionally, to make the simulator easy for unskilled users to operate, the application of experimental units into the inputter section is desirable and, consequently, they have to be converted to simulation units (molar).

The initial number of individuals of a given simulation, $I(t=0)$, is calculated according to a provided initial cell density (individuals/ml) and the simulation volume. Then, n_0 individuals are taken at random from a text file, saved from a previous simulation (inoculum bank), keeping individual variable values except for their position (X,Y,Z), which is taken randomly and equiprobably within the spatial domain, and its carbohydrates reserves (R_{GLU}), which is assumed to be 0. This initialization of the individual's state allows the influence of the initial (and uncertain) values under the system evolution to be reduced. Initialization of the individuals in the simulation used as inoculum source is summarized in Table A.1.

Table A.1. Initialization of individual state variables at the beginning of the simulation used as inoculum. Within the table, a denotes the initial value, iq is the number of spatial cells per dimension and m_C is the critical mass at reproduction model parameter, and σ_{mc} its standard deviation.

Individual state variable	Initial value
M	$a = m_C \cdot U(0.8,1.2)$
M_{Start}	$a; a \rightarrow N(m_C, \sigma_{mc})$
B	0
P	1
M_{inc}, T_{inc}, R_{GLU}, R_{CN}, C_{GLU}^{in}, D	0
X,Y,Z	$a\ ; P[a=k] = \frac{1}{iq}; k = 1, \dots, iq$

Spatial cells state variables are initialized according to an initial concentration (g/l) which is equally distributed among all spatial cells. Substrate molar mass, simulated volume and the number of spatial cells allow changing from initial concentration to initial cell content (pmol per spatial cell). Organic labile N is assumed to have the following elemental formula: $\mathrm{C}_{a_C}\mathrm{N}$.

A.6 Input data

The model does not use input data to represent time-varying processes.

A.7 Submodels

A.7.1 Motion submodel

Since a discrete space is modelled, the position of the cells is defined by three integer coordinates. It is assumed that the position of a yeast cell may change to a new position in the space. The yeast cell position after the movement is randomly chosen among the current spatial cell of coordinates (x_0,y_0,z_0) and a 3D Moore neighbourhood, $D^M_{x_0,y_0,z_0}$, as defined by Zhang et al. (2009):

$$D^M_{x_0,y_0,z_0} = \begin{Bmatrix} (x,y,z): 0 < |x-x_0| + |y-y_0| + |z-z_0| \leq 3 \\ with\ |x-x_0|, |y-y_0| and |z-z_0| \in \{0,1\} \end{Bmatrix} \quad (A.3)$$

where (x,y,z) represents the coordinates of the neighbourhood spatial cells of the cell of coordinates (x_0,y_0,z_0). We are using periodic boundary conditions.

A.7.2 Uptake submodel

Yeast translocate low molecular weight compounds dissolved in water through the cell membrane, which represent the fundamental way in which they communicate and interact with their surrounding environment (Walker, 1998). From the various sugars that can be used by *S. cerevisiae*, glucose plays a distinguished role in respirative and fermentative metabolism and is widely used in laboratory experiments. Additionally, glucose exhibits a repressive and inhibitory effect on the assimilation of other sugars. When it comes to N sources, *S. cerevisiae* is capable of transporting and using a few inorganic (ammonium ions and urea) and several organic N, mainly amino acids and some peptides (di o tri peptides), with specific transport mechanisms. Ammonium salts are routinely incorporated into yeast growth media because all yeasts are able to use ammonia efficiently. In brewing fermentations mixtures of amino acids provide the bulk of the yeasts' N supply. Peptides are available in industrial media through proteolysis of plant protein (hordein from barley in brewing) and through the addition of peptones in laboratory media.

Three substrates are taken into account by the model: glucose, organic N and ammonium (j=GLU, CN, NH4). As previous INDISIM models (Ginovart et al., 2005; Gras et al., 2011), it is assumed that the maximum individual uptake of C and N sources is controlled by the internal C to N ratio (Eq. A.4), a value lower than or equal to r_C to uptake C sources and greater than or equal to r_N to uptake N sources.

$$C/_N(t) = \frac{M(t) \cdot a_{yC} + R_{GLU}(t) \cdot 6 + R_{CN}(t) \cdot a_C + C_{GLU}(t) \cdot 6}{M(t) + R_{CN}} \quad (A.4)$$

where a_C is the C to N molar ratio of the simulated organic labile compound (CN).

The maximum individual uptake, $U^j_{MAX}(t)$, in a time step for the substrate j (pmol), will be determined by the size of the individual (i.e. its structural

mass). As scar tissue is thought to be less efficient than normal cell wall material at facilitating the transport of macromolecules into the cell (Powell et al., 2003a), it has been previously proposed that an increase in the bud scars number can limit the surface area for nutrient exchange (Mortimer and Johnston, 1959). Interestingly, as shown by Powell et al. (2003a), bud scars (as birth scar) expand with genealogical age and, as a result, the percentage of cell surface occupied by a scar is almost constant throughout the genealogical lifespan of the cell, making the effects of the number of scars of the cell, $B(t)$, constant (Eqs. A.5 and A.6). It is worth noting that this may not be true if the individual's mass, and not its surface area, is being considered in its own right. Assuming a characteristic cell shape and density, the relationship between the cell surface area, cell volume, and cell mass, which is usually used to obtain experimentally the unitary uptake rate, can be approached (see, for instance, Ginovart et al., 2005; Gras et al., 2010, 2011). Keeping in mind the high level of uncertainty of the necessary estimates (which makes this computation highly uncertain), this unit transformation has been avoided. Nonetheless, a constant effect of the number of scars on the uptake has been taken into account. Additionally, in the case of glucose (Eq. A.6), the model assumes that the maximum uptake is also affected by the extracellular ethanol content of the spatial cell the yeast is currently in. This toxic effect of the ethanol has already been taken into account in previous modelling efforts (Steinmeyer and Shuler, 1989; Sainz et al., 2003).

$$U^{j}_{MAX}(t) = \ddot{u}_j \cdot M(t) \cdot \big(1 - p_s\, B(t)\big); j = \mathrm{CN}, \mathrm{NH4} \tag{A.5}$$

$$U^{j}_{MAX}(t) = \ddot{u}_j \cdot M(t) \cdot \big(1 - p_s\, B(t)\big) \cdot \left(1 - \frac{S_{\mathrm{OH}}(t)}{\ddot{p}_{u\mathrm{OH}}}\right); j = \mathrm{GLU} \tag{A.6}$$

where $\ddot{u}_j$ is the specific substrate j uptake rate (pmol/pmol $\mathrm{CN_{MIC}}$/TS), p_s is a penalization per cell scar and $\ddot{p}_{uOH}$ is the tolerance to ethanol of the yeast (pmol Ethanol/SC). $\ddot{u}_j$ is a non-negative random draw from a normal distribution with mean u_j and standard deviation σ_{uj}. $\ddot{p}_{u\mathrm{OH}}$ is a non-negative random draw from a normal distribution with mean $p_{u\mathrm{OH}}$ and standard deviation $\sigma_{pu\mathrm{OH}}$. Negative numbers for $U^{j}_{MAX}(t)$ are not allowed and the value of zero is used instead.

The substrate uptake is also influenced by the amount of available substrate, that is the amount of substrate j present in the spatial cell, $S_j(t)$, in which the yeast is evolving. The substrate j uptake (pmol) at a given time step, U_j(t), is then the minimum quantity between the $U^{j}_{MAX}(t)$ and $S_j(t)$.

A.7.3 Metabolism submodel

For clarity, the metabolism submodel is divided into these sections: (i) variables update, (ii) maintenance requirements, (iii) carbon reserves creation section, (iv) new mass synthesis, and (v) substance release.

(i) *Variables update*. The internal amount of glucose able to be metabolized is given by its glucose uptake in a time step and the amount of non-metabolized glucose remaining from the previous time step, therefore, the glucose uptake is modified as follows: $U_{\mathrm{GLU}}(t) = U_{\mathrm{GLU}}(t) + C^{in}_{\mathrm{GLU}}(t-1)$. Similarly, organic N uptake is also recalculated to take into account cell's internal reserves: $U_{\mathrm{CN}}(t) = U_{\mathrm{CN}}(t) + R_{\mathrm{CN}}(t\text{-}1)$.

(ii) *Maintenance requirements*. The model assumes the fermentative catabolic pathway (glycolysis and alcoholic fermentation) of *S. cerevisiae* as the first option to achieve metabolic energy. The following widely-accepted (Madigan, 2006) coupled global reaction (Eq. A.7) allows us to establish the metabolism model:

$$1\ C_6H_{12}O_6 \rightarrow 2\ C_2H_6O + 2\ CO_2 + 2\ ATP \tag{A.7}$$

A yeast cell requires energy for its cellular maintenance, and it is assumed that this energy is proportional to its structural mass, $M(t)$. Additionally, the ethanol accumulated during the fermentation is a potent chemical stress factor for yeast cells so various physiological adaptations, which are thought to protect it against ethanol, occur in the cell (Walker, 1998), and measured maintenance requirements are known to be increased (Yang et al., 2012). Hence, if e is the maintenance rate through fermentative metabolism (pmol glucose/pmol of CN_{MIC}/TS) that a cell needs to remain viable for the current time step, we assume that the total energy required to maintain viability by fermenting glucose, $Ma(t)$, is given by Eq. A.8:

$$Ma(t) = M(t) \cdot \left(e + \ddot{p}_{\mathrm{OH}}\, S_{\mathrm{OH}}(t)\right) \tag{A.8}$$

where $\ddot{p}_{e\mathrm{OH}}$ is a penalization due to local ethanol content (pmol of glucose/pmol of CNMIC/(pmol of ethanol/SC)). $\ddot{p}_{OH}$ is a non-negative random draw from a normal distribution with mean p_{OH} and standard deviation p_{OH}. The model assumes that maintenance energy must be covered first before a cell is able to experience growth, i.e. new mass synthesis.

Taking Eq. A.8 into account, if $U_{\mathrm{GLU}}(t) \geq Ma(t)$, so the cell has enough glucose to cover Ma, it remains viable, and $U_{\mathrm{GLU}}(\mathrm{t})$ is updated by subtracting Ma. The remaining $U_{\mathrm{GLU}}(\mathrm{t})$ will then be used to create C reserves or new mass (see below). But if $U_{\mathrm{GLU}}(t) < \mathrm{Ma(t)}$, then the cell does not have enough glucose to cover Ma and it needs to use its C reserves, $R_{\mathrm{GLU}}(\mathrm{t})$. In that situation, if $U_{\mathrm{GLU}}(t) + R_{\mathrm{GLU}}(t) \geq Ma(t)$, then the cell has enough glucose to cover Ma and remains viable, but without being capable of creating new mass; however, if $U_{\mathrm{GLU}}(t) + R_{\mathrm{GLU}}(t) < Ma$, it is not able to fulfil Ma, and may die (see section A.7.5 lifespan submodel below). Individuals that have completed their maintenance and still have glucose inside the cell are able to use it in the creation of new mass or new C reserves.

(iii) *Carbon reserve creation*. Resting and stressed cells accumulate high concentrations of storage carbohydrates. The main storage carbohydrate in yeast cells are trehalose and glycogen (Panek, 1991). Glycogen degradation and synthesis are controlled by environmental factors (e.g. Rowen et al., 1992).

Once exogenous sugars and other essential nutrients become limiting, yeast can accumulate glycogen when their cell cycle is arrested. Using the stored glycogen under starvation conditions by endogenous fermentation contributes to cellular maintenance by providing ATP to keep cell viability (Walker, 1998). During the initial stages of fermentation of brewer's wort, *S. cerevisiae* rapidly degrades stored glycogen. To model this behaviour, it is assumed that reserve carbohydrates are created, and thus $R_{GLU}(t+1)=U_{GLU}(t)$, when glucose content in the spatial cell is below a threshold (l_S). Otherwise, new mass can be synthesized.

(iv) *New mass synthesis*. In order to synthesize new mass, *S. cerevisiae* needs a C source, an N source, and a certain amount of energy. Yeast anabolic reactions (energy-consuming, reductive process) have to be coupled to catabolic reactions (energy-giving, oxidative process), which is achieved by dehydrogenase enzymes which predominantly use NADP and NAD, respectively, as redox cofactors (Walker, 1998). Combing stated catabolic (Eq. A.7) with the anabolic needs of C, N, and energy, Eq. A.9 and A.10 can be obtained. Notice that, in addition to its energetic role, glucose is the only C source to fulfil the anabolic part of the coupled reaction A.9. In contrast, since organic N is simultaneously a C and N source, both glucose and organic N are providing the needed anabolic C within Eq. A.10.

$$a_{yc}\,\mathrm{C_{GLU}} + 1\,\mathrm{N_{NH4}} + E_1 \rightarrow 1\,\mathrm{C}_{a_{yc}}\mathrm{N}$$
$$\alpha_1\,\mathrm{C_{GLU}} \rightarrow \alpha_1\,\frac{4}{6}\,\mathrm{C_{OH}} + \alpha_1\,\frac{2}{6}\,\mathrm{C_{CO2}} + \frac{2}{6}\,\alpha_1$$
$$E_1 = \frac{2}{6}\,\alpha_1$$

$$\left(a_{yc}+\alpha_1\right)\mathrm{C_{GLU}} + 1\,\mathrm{N_{NH4}} \rightarrow 1\,\mathrm{C}_{a_{yc}}\mathrm{N} + \frac{4}{6}\alpha_1\,\mathrm{C_{OH}} + \frac{2\alpha_1}{6}\mathrm{C_{CO2}} \tag{A.9}$$

$$\left(a_{yc}-a_c\right)\mathrm{C_{GLU}} + a_c\,\mathrm{C_{CN}} + 1\,\mathrm{N_{CN}} + E_2 \rightarrow 1\,\mathrm{C}_{a_{yc}}\mathrm{N}$$
$$\alpha_2\,\mathrm{C_{GLU}} \rightarrow \alpha_2\,\frac{4}{6}\,\mathrm{C_{OH}} + \alpha_2\,\frac{2}{6}\,\mathrm{C_{CO2}} + \frac{2}{6}\,\alpha_2$$
$$E_2 = \frac{2}{6}\,\alpha_2$$

$$\left(a_{yc}-a_c+\alpha_2\right)\mathrm{C_{GLU}} + a_c\,\mathrm{C_{CN}} + 1\,\mathrm{N_{CN}} \rightarrow 1\,\mathrm{C}_{a_{yc}}\mathrm{N} + \frac{4}{6}\alpha_2\,\mathrm{C_{OH}} + \frac{2\alpha_2}{6}\mathrm{C_{CO2}} \tag{A.10}$$

where $\mathrm{C_{GLU}}$ are C moles from glucose, $\mathrm{C_{CN}}$ are C moles from the modelled organic N, $\mathrm{N_{NH4}}$ are N moles from ammonium, $\mathrm{N_{CN}}$ are N moles from the modelled organic N, $\mathrm{C}_{a_{yc}}\mathrm{N}$ N is a mole of yeast biomass, $\mathrm{C_{OH}}$ are C moles from

the ethanol, C_{CO2} are C moles from the carbon dioxide, α_1 and α_2 indicate the C moles from glucose needed to synthetize a mass unit (mol C-glucose/mol CNMIC); and E_1 and E_2 are metabolic energy required to create a mass unit (mol ATP). Taking into account that energy is usually stored inside microorganisms as macromolecules so the amount of ATP remains low (Madigan, 2006), Eqs. A.9 and A.10 assume that the addition of metabolic energy from anabolic and catabolic subreactions becomes zero.

Knowing the remaining amounts of C inside the cell, $U_{GLU}(t)$; organic N, $U_{CN}(t)$ and ammonium $U_{NH4}(t)$, the C and N composition of the substrates and the stoichiometry of the global reactions (Eqs. A.9 and A.10), the moles of new mass, Δm, will be defined by the minimum value found multiplying the amount of reactants by the C (or N) moles of the substrate divided by the corresponding stoichiometric coefficient. Finally, the submodel assumes that a cell will synthesize new mass consecutively following Eqs. A.9 and A.10. Once the newly created mass is known, the individual mass is set to: $M(t) = M(t\text{-}1)+\Delta m$. Similarly, if the individual is budding, the increased mass since the beginning of the budding phase, M_{inc}, is also updated $M_{inc}(t) = M_{inc}(t\text{-}1)+\Delta m$. If N sources limit the mass synthesis, and thus some glucose still remains inside the cell, it is stored in the corresponding individual state variable. Therefore, $C^{in}_{GLU}(t) = U_{GLU}(t)$.

(v) *Substance release*. As a consequence of mass creation, the amount of fermented glucose is also known. From this, metabolic heat and products are released to the spatial cell. The non-used ammonium is released to the medium but, taking into account that yeast cells store large pools of endogenous amino acids (Eddy, 1980; Steinmeyer and Shuler, 1989), the model considers that remaining organic N is stored as internal N reserves, R_{CN}. Therefore, $R_{CN}(t) = U_{CN}(t)$.

A.7.4 Reproduction submodel

The model assumes that a cell can start a new cell cycle only after the preceding cycle is completed, and that the cellular cycle involves two differentiated phases. Phase 1, or unbudded phase ($P(t)=1$), covers most of phase G1 and a very small fraction of phase S in the traditional division of the cell cycle; while phase 2, or budding phase ($P(t)=2$), covers a small fraction of G1, most of S and all of G2 and M (see Ginovart and Cañadas, 2008; Prats et al., 2010; Ginovart et al. 2011a;b, and references therein).

Conceptually, the model assumes that in the unbudded phase the yeast cell is getting ready for budding and that change to the budding phase takes place only when the cell: (i) has attained a minimum cellular mass, defined by the parameter m_C, the critical or minimum mass; and (ii) has achieved a minimum growth of its mass, which is related to the model parameter Δm_{B1}. Note that within the model, Phase 1 does not need to be completed in a given time interval. In practice, this behaviour is achieved by comparing, every time step,

if the cell mass $M(t)$, has attained its needed mass at "Start", $M_{Start}(t)$, which is set at the beginning of Phase 1 so it fulfils the stated conceptual assumptions (see computation of M_{Start} below). Therefore, if, at a given time step, $M(t) \geq M_{Start}(t)$, the cell enters the budding phase. Note that this check in the model is to see whether an individual cell has reached a "Start" mass, irrespective of its original value and growth rate.

The budding phase is the least flexible in the cellular cycle as it requires both temporal and growth checks. Within the model, two conditions must be satisfied for the releasing of the bud, and the subsequent change to the unbudded phase. These are: (i) a minimum growth of mass, which is related to the parameter Δm_{B2}; and (ii) a minimum time interval, which is related to the parameter Δt_2. The first requirement is met when $M_{inc} \geq \ddot{\Delta m}_{B2}$; and the second when $T_{inc} > \ddot{\Delta t}_{B2}$, where $\ddot{\Delta m}_{B2}$ (pmol CNMIC) and $\ddot{\Delta t}_{B2}$ (TS) are non-negative random draws following, respectively, N(Δm_{B2}, σ_{B2}) and N(Δt_2, σ_{t2}). The first condition is necessary because a yeast cell must have a minimum number of molecules and achieve minimum structural mass in order to function as an independent entity. On the other hand, the bud growth, even under optimal growth conditions, needs a minimum time interval to be completed; this is represented by the second condition.

The budding phase is completed with the cell division. The mass of the mother cell after the cell division, $M(t+1)$, is equal to the actual mass less the daughter mass (m_D) as follows: $M(t+1)=M(t)-m_D$. As previously assumed in other modelling efforts (Porro et al., 2009), once the cell passes "Start", all the newly synthesized mass goes to the daughter cell, which has been experimentally corroborated (Alberghina et al., 1998 and references therein). Thus, $m_D=M_{inc}(t)$, represents the accumulated mass since the start of the budding phase until t.

Once m_D and $M(t+1)$ are known, and since both mother and daughter cells start a new unbudded phase, their mass at "Start", $M_{Start}(t+1)$, have to be computed. Taking the mother cell as an example, If $M(t+1) \leq m_C - \Delta m_{B1}$ then the "Start" mass assigned to the cell is $M_{Start}(t+1)= \ddot{m}_C$; otherwise, if $M(t+1) > m_C - \Delta m_{B1}$ then the "Start" mass assigned to the cell is $M_{Start}(t+1) = M(t+1) + \ddot{\Delta m}_{B1}$ where $\ddot{m}_C$ and $\ddot{\Delta m}_{B1}$ are positive random draws following, respectively, N(m_C, σ_{mC}) and N(Δm_{B1}, σ_{mC}). The calculus for the daughter cell is analogous. Notice that using this method, the previously stated conceptual assumptions of the requirements of the unbudded phase are met.

Properties of the, until now, mother cell are shared with the new daughter cell as a result of the bud detachment. The model assumes that this distribution of properties is proportional to the fraction of the detached mass. Therefore, and as previously stated, the fraction of the detached mass can be obtained from the relation:

$$\frac{M(t+1)}{M(t)} + \frac{m_D}{M(t)} = 1 \tag{A.11}$$

Using the previous expression and knowing $R_{\mathrm{GLU}}(t)$, $R_{\mathrm{CN}}(\mathrm{t})$ and $C^{in}_{\mathrm{GLU}}(t)$ for the mother cell, $R_{\mathrm{GLU}}(_{\mathrm{t}}+1)$, $R_{\mathrm{CN}}(t+1)$ and $C^{in}_{GLU}(t+1)$ for both the mother and the daughter cell can be easily computed.

The daughter cell (a virgin cell since its number of bud scars is 0) is born in the same spatial cell occupied by the mother. Moreover, since the previous history of the mother cell is affecting both the cell and the bud, $D(t+1)$, the time without covering maintenance requirements of the daughter cell is also inherited from the main cell at detachment. The daughter cell remains active since the reproduction process has finished but it does not act until the next time step is reached.

A.7.5 Lifespan submodel

According to Maskell et al. (2003), it is essential to define the chronological lifespan as a consequence of cumulative and irreversible damage to intracellular components during extended stationary phase, which compromises cell integrity and leads to death and autolysis. In contrast, replicative lifespan is related to the number of divisions that an individual cell undertakes before entering a non-replicative state termed senescence, which leads to cell death and autolysis (Maskell et al., 2003; Jenkins et al., 2003 and references therein).

Both chronological and replicative lifespan are taken into account within the model. Chronological lifespan is followed using a mortality index, $D(t)$, which satisfies the following rules: (i) whenever a cell is not able to satisfy its maintenance requirements in a time step, the simulator increases the mortality index of the cell by one: $D(t+1)=D(t)+1$; conversely, (ii) the index is set to 0, $D(t+1)=0$. Chronological lifespan is evaluated at every time step by comparing $D(t)$ and $\ddot{t}_d$ (TS), a non-negative random draw from an $\mathrm{N}(t_d, \sigma_{td})$. Here, t_d denotes an average time beyond which the cell cannot survive. Thus, if $D(t)>t_d$, the cell dies. In the model the individual cells may die, indirectly, due to ethanol excess or low glucose concentration. These reduce the ability of the yeast to cover its maintenance energy and increase mass. The magnitude of these unfavourable conditions will determine the vitality and viability of the individual yeast cell. Replicative lifespan is followed using the cell's replicative age (i.e. the number of cell scars), $B(t)$. Once a cell reaches the maximum number of scars it can experience, known as the Hayflick limit (Hayflick, 1968; Jenkins et al. 2003, and references therein), h_l, the yeast cell dies. When the cell dies, its mass remains in the medium and no turnover is modelled.

A.7.6 Actions over the medium

Substrates are subject to diffusion processes through the three-dimensional lattice of spatial cells according to Fick's Law (Bormann et al., 2001). The

model accounts for mass transfer among cells situated into a 3D Moore neighbourhood (Eqs. A.12-A.15). Given a spatial cell of coordinates (x,y,z), S_{xyz}, containing, at time t, a specific amount of substrate j, designated here as $S_j^{xyz}(t)$; and accounting for the amount of diffused substrate from (into) the neighbourhood cells, the amount of substrate j within the cell after diffusive processes, $S_j^{xyz}(t+1)$, can be obtained by:

$$S_j^{xyz}(t+1) = S_j^{xyz}(t) + d_j \cdot \left[\Delta_c + \frac{\Delta_{d1}}{\sqrt{2}} + \frac{\Delta_{d2}}{\sqrt{3}} - \left(6 + \frac{12}{\sqrt{2}} + \frac{8}{\sqrt{3}}\right) \cdot S_j^{xyz}(t)\right] \quad \text{(A.12)}$$

where d_j accounts for the fraction of substrate j diffused from (into) the spatial cell S_{xyz} to (from) any spatial cell with centres are at one spatial cell of distance, and:

$$\Delta_c = S_j^{x-1yz}(t) + S_j^{x+1yz}(t) + S_j^{xy-1z}(t) + S_j^{xy+1z}(t) + S_j^{xyz-1}(t) + S_j^{xyz+1}(t) \quad \text{(A.13)}$$

$$\begin{aligned}\Delta_{d1} = {} & S_j^{xy+1z-1}(t) + S_j^{xy+1z+1}(t) + S_j^{xy-1z-1}(t) + S_j^{xy-1z+1}(t) \\ & + S_j^{x+1yz-1}(t) + S_j^{x+1yz+1}(t) + S_j^{x-1yz-1}(t) + S_j^{x-1yz+1}(t) \\ & + S_j^{x+1y-1z}(t) + S_j^{x+1y+1z}(t) + S_j^{x-1y-1z}(t) + S_j^{x-1y+1z}(t)\end{aligned} \quad \text{(A.14)}$$

$$\begin{aligned}\Delta_{d2} = {} & S_j^{x+1y+1z+1}(t) + S_j^{x+1y+1z-1}(t) + S_j^{x+1y-1z+1}(t) + S_j^{x+1y-1z-1}(t) \\ & + S_j^{x-1y+1z+1}(t) + S_j^{x-1y+1z-1}(t) + S_j^{x-1y-1z+1}(t) \\ & + S_j^{x-1y-1z-1}(t)\end{aligned} \quad \text{(A.15)}$$

A.8 Implemented model verification

A variety of measures and techniques have been used in order to make sure the model was accurately implemented. As a general measure, the INDISIM-*Saccha* simulator, unlike previous INDISIMs, has been implemented using Fortran 90 code in order to take advantage of its new features (mainly, free-form source input, intrinsic array functions, dynamic memory allocation, etc.) which allow for a source code simplification, thus, making the code less error prone. Although newer language specifications exist, Fortran 90 is a qualitative leap from previous versions but it nevertheless allows compatibility with existing code. It is also worth noting that, aiming to avoid otherwise potentially hard to find errors, implicit variables declaration has been avoided by using the Implicit None declaration in all the program units.

Internal model logics and behaviour were systematically investigated by collecting both global and individual data through the simulation length and visually testing them within a spreadsheet program. In particular, global variables include: (i) yeast growth curve; (ii) accumulated death individuals curve; (iii) substrates (glucose, organic N and NH_4) uptake curves; (iv) main metabolites production curves (released ethanol and CO_2); and (v) substrate

effective uptake (defined as uptaken substrate divided by structural mass of the population at the beginning of the time step). Individual data were investigated by way of histogram and included: (i) structural mass distribution; and (ii) genealogical age distribution. It is worth pointing out the dual purpose of this practice which facilitates understanding of the model (or modelled system) and of the consequences of the modelling decisions undertaken.

Because it is crucial to maintain mass balances of the chemical elements considered in this design, a central piece of the verification of INDISIM-*Saccha* is a C and N mass balance built internally into the simulator. Although not a guarantee of perfect implementation, it allows detection of otherwise hard to find errors. Specifically, the C and N contents within the system at any particular time step have been followed by means of the existing fraction (%) of the element regarding its initial amount at systems initialization adjusted to 6 decimal points.

Main parts of the final version of the simulator were also thoroughly checked by means of specifically built testing programs which allowed us to compare their output against an independent reimplementation of the submodel part on a different software platform or, when appropriate, by testing statistically the output of the submodel being inspected.

Inputter and Outputter subunits and the motion, uptake, metabolism, reproduction, and actions over the medium submodels of the implemented simulator were formally verified. Input values transformation from experimental to internal simulator units were recorded and tested by comparison against an independent reimplementation of the necessary calculations using spreadsheet software. Transformation of the model outputs from internal to experimental units were recorded and tested by comparison against an independent reimplementation of the required calculations using spreadsheet software. Motion submodel was verified using two independent tests investigating the crossing of the medium boundaries by the moving cells and the probability of the movement to the neighbourhood. The first test of the motion submodel was conducted by simulating 20 random movements from all the existing cells of the three-dimensional space and checking the coordinates after the movement to be between 1 and iq, the number of cells per dimension. For a robust design of the submodel, iq values from 1 to 5 were essayed. The second test of the motion submodel was done by simulating 5400 random movements from a fixed initial spatial cell, recording coordinates after the movement, and obtaining occurrence of the movement to the 27 allowable cells with spreadsheet software . Equality of frequencies was statistically tested in R (R Core Team, 2012) by means of the chisq.test function of the stats package (p-value>0.98). Uptake and metabolism submodels were thoroughly tested by randomly setting their input variables according to their allowable values and subsequently reproducing subroutine output by an independent reimplementation of the necessary computations using spreadsheet software . The reproduction submodel was thoroughly tested by randomly setting input variables of the submodel, according to their allowable values, and

subsequently reproducing subroutine output by an independent reimplementation within spreadsheet software. A sufficient number of repetitions (2000) allowed exploration of all subunit parts. Although a number of minor changes were made to the final reproduction submodel version, the high chance of any error within this submodel impairing C and N balances and a careful upgrading of the program unit allow us to assume there are no significant errors. Finally, actions over the medium (diffusive processes) submodel was tested by comparison with the output of the testing program against an independent reimplementation of the submodel in R (R Core Team, 2012).

References

Alberghina, L., Smeraldi, C., Ranzi, B.M, Porro, D., 1998. Control by nutrients of growth and cell cycle progression in budding Yeast, analyzed by double-tag flow cytometry. *Journal of Bacteriology* 180 (15), 3864-3872.

Bormann, G., Brosens, F., DeSchutter, E., 2001. Diffusion, in: Bower, Bolouri (Eds.), *Computational Modelling of Genetic and Biochemical Networks*, The MIT Press, Boston.

Eddy A.A., 1980. Some aspects of amino acid transport in yeast. In *Microorganisms and Nitrogen Sources* (ed. J.W.Payne), pp. 35-62. J. Wiley and Sons Ltd., Chichester.

Ginovart, M., Cañadas J.C., 2008. INDISIM-YEAST: an individual-based simulator on a website for experimenting and investigating diverse dynamics of yeast populations in liquid media. *Journal of Industrial Microbiology and Biotechnology* 35(11), 1359-1366. doi: 10.1007/s10295-008-0436-4.

Ginovart, M., López, D., Gras, A., 2005. Individual-based modelling of microbial activity to study mineralization of C and N and nitrification process in soil. *NonLinear Analysis: Real World Applications* 6(4), 773-795. doi: 10.1016/j.nonrwa.2004.12.005

Ginovart, M., Prats, C., Portell, X., Silbert, M., 2011a. Exploring the lag phase and growth initiation of a yeast culture by means of an Individual-based Model. *Food Microbiology* 28(4), 810-817. doi: 10.1016/j.fm.2010.05.004

Ginovart, M., Prats, C., Portell, X., Silbert, M., 2011b. Analysis of the effect of inoculum characteristics on the first stages of a growing yeast population in beer fermentations by means of an Individual-based Model. *Journal of Industrial Microbiology and Biotechnology* 38(1), 153-165. doi: 10.1007/s10295-010-0840-4.

Gras, A., Ginovart, M., Portell, X., Baveye, P.C., 2010. Individual-based modelling of carbon and nitrogen dynamics in soils: parameterization and sensitivity analysis of abiotic components. *Soil Science* 175, 363-374. doi: 10.1097/SS.0b013e3181eda507

Gras, A., Ginovart, M., Valls, J., Baveye, J.C., 2011. Individual-based modelling of carbon and nitrogen dynamics in soils: parameterization and

sensitivity analysis of microbial components. *Ecological Modelling* 222, 1998-2010. doi: 10.1016/j.ecolmodel.2011.03.009.

Grimm V., Berger, U., Bastiansen, F., Eliassen S., et al., 2006. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* 198 (1-2), 115-126. doi: 10.1016/j.ecolmodel.2006.04.023.

Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F., 2010. The ODD protocol: a review and first update. *Ecological Modelling* 221 (23), 2760-2768. doi: 10.1016/j.ecolmodel.2010.08.019.

Hayflick, L., 1968. The limited in vitro lifetime of human diploid cell strains. *Experimental Cell Research* 37 (3), 614-636. doi: 10.1016/0014-4827(65)90211-9

Jenkins, C.L., Kennedy, A.I., Hodgson, J.A., Thurston, P., Smart, K.A., 2003. Impact of serial repitching on lager brewing yeast quality. *Journal of the American Society of Brewing Chemists* 61 (1), 1-9.

Madigan, M.T., 2006. *Brock Biology of microorganisms*. 11th ed. Upper Saddle River: Prentice Hall, 2006. ISBN 0-13-196893-9.

Maskell, D.L., Kennedy, A.I., Hodgson, J.A., Katherine, A.S., 2003. Chronological and replicative lifespan of polyploid *Saccharomyces cerevisiae* (syn. *S. pastorianus*). *FEMS Yeast Research* 3 (2), 201-209. doi: 10.1016/S1567-1356(02)00199-X

Mortimer, R.K., Johnston, J.R., 1959. Life span of individual yeast cells. *Nature* 183, 1751-1752. doi: 10.1038/1831751a0

Panek, A.D., 1991. *Storage carbohydrates*. In The Yeasts, 2nd ed., Vol.4. Yeast Organelles (eds A.H. Rose and J.S. Harrison), pp. 655-678. Academic Press, London.

Powell, C.D., Quain, D.E., Smart, K.A., 2003. Chitin scar breaks in aged *Saccharomyces cerevisiae*. *Microbiology* 149 (11), 3129-3137. doi: 10.1099/mic.0.25940-0.

Porro, D., Vai, M., Vanoni, M., Alberghina, L., Hatzis, C., 2009. Analysis and modeling of growing budding yeast populations at the single cell level. *Cytometry A* 75 (2), 114-120. doi: 10.1002/cyto.a.20689.

Prats, C., Ferrer, J., Gras, A., Ginovart, M., 2010. Individual-based modelling and simulation of microbial processes: yeast fermentation and multi-species composting. *Mathematical and Computer Modelling of Dynamical Systems* 16 (6), 489-510.

R Core Team, 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rowen, D.W., Meinke, M., LaPorte, D.C., 1992. GLC3 and GHA1 of *Saccharomyces cerevisiae* are allelic and encode the glycogen branching enzyme. *Molecular and Cellular Biology* 12 (1), 22-29.

Sainz, J., Pizarro, F., Pérez-Correa, J.R., Agosin, E., 2003. Modeling of yeast metabolism and process dynamics in batch fermentation. *Biotechnology and Bioengineering* 81 (7), 818-828. doi: 10.1002/bit.10535.

Steinmeyer, D.E., Shuler, M.L., 1989. Structured model for *Saccharomyces cerevisiae*. *Chemical Engineering Science* 44 (9), 2017-2030. doi: 10.1016/0009-2509(89)85138-3.

Walker, G.M., 1998. *Yeast physiology and biotechnology*. Wiley, Chichester.

Yang, K.M., Lee, N.R., Woo, J.M., Choi W., Zimmermann, M., Blank, L.M., Park, J.B., 2012. Ethanol reduces mitochondrial membrane integrity and thereby impacts carbon metabolism of *Saccharomyces cerevisiae*. *FEMS Yeast Research* 12 (6), 675-84.

Zhang, L., Strouthos, C.G., Wang, Z., Deisboeck, T.S., 2009. Simulating brain tumor heterogeneity with a multiscale agent-based model: linking molecular signatures, phenotypes and expansion rate. *Mathematical and Computer Modeling* 49 (1-2), 307-319. doi: 10.1016/j.mcm.2008.05.011.

Appendix B

Parameterization of INDISIM-*Saccha*

B.1 Temporal resolution of the model

Tyson et al. (1979) experimentally obtain the doubling time of batch cultures of *S. cerevisiae* growing in a number of C sources at 30 °C and calculates the corresponding population budding time. The fastest growing cells (71 min) showed a budding time of about 62 min, which can be used as the fastest discrete event modelled. A temporal resolution that was too low would not be consistent with the variability taken into account by the model while a temporal resolution that was too high would have non-assumable time computing costs. Keeping this in mind, a time step length of 0.1 h is adopted. The simulation time horizon is set to 14 days to match experimental data used.

B.2 Spatial resolution of the model

In order to be able to capture microscopic level variability, a minimum number of individuals must be simulated at any time. In our conditions, the minimum yeast density to be simulated has an order of magnitude of $1x10^5$ cells/ml, and considering a minimum of 10 simulated individuals, the simulated volume is set to $1x10^{-4}$ ml.

As before, too low a density could impede the possible interactions between the yeasts for spatial reasons while too high a density would be too computationally expensive. To achieve computational-cost effective simulations, the number of spatial cells per cube side, iq, has been fixed at 10, giving a three-dimensional lattice of 10^3 elements, a reasonable number of elements given that a liquid medium is being simulated. Although not specifically prevented, the maximum number of individuals physically fitting into one cell, as computed using a cell diameter of 5.5 μm (Portell et al., 2011) and assuming a spherical shape, were not reached during the simulations.

B.3 Initialization of the simulations

As previously explained, inputted parameters which are expressed in experimental units, have to be changed to simulation units. According to the previously suggested "empirical formula" for baker's yeast (Berry, 1988): $C_{47}H_{6.3}O_{33}N_8P_{1.2}salts_{4.5}$, and adopting the calcium chloride molar mass, it follows that the yeast dry mass C composition is approximately 32 % (w/w). From the stoichiometric formula, the C to N ratio, ayc is approximated to 6. This data, taken jointly with the total number of spatial cells, the simulated volume of liquid media, the fixed temporal resolution, and the molar weights, are used in order to translate from mass and temporal units into the units used internally by the simulator, pmol and time steps.

Three initial glucose concentrations were tested at the beginning of the simulations (150, 200 and 250 g/l) according to experimental practice while

remaining chemical species were fixed throughout the experimental treatments. Initial concentrations of ethanol and carbon dioxide were set to 0 g/l.

Setting the initial concentration of the N species was done by distributing the measured Kjeldahl N (0.0152 g N/l) according to the theoretical N contents coming from the medium ingredients used. These theoretical ammonium and organic N contents were derived by using the composition of the growth medium and the α-Amino N (N_{AMINO}) and Total N (N_{TOT}) typical contents of the ingredients as extracted from published data from the culture media manufacturer (Scharlau, 2011). The N_{TOT} content of the yeast extract was split into organic N (N_{ORG}) and ammonium ($N_{TOT}=N_{NH4}+N_{ORG}$) while Meat Peptone was assumed to be a source of organic N solely ($N_{TOT}=N_{ORG}$). The fractions of N_{NH4} and N_{ORG} of the whole growth medium derived this way were 33 and 67 % respectively.

Simulated organic labile N brings C in its formulation, with a_C its C to N molar ratio. a_C has also been derived from published data (Scharlau, 2011) by summation of the C to N relation of the composing amino acids weighted by their relative abundance as dictated by the free content of the amino acid of an ingredient (g free amino acid/g DW) by the used concentration of the ingredient in the culture media (g DW/l). From this, a_C is approximated to 4.5.

The initial number of individuals, $I(t=0)$, used to initialize the population at the starting of the simulation is treatment dependent and set according to the mean initial total cell density obtained experimentally (Table 1 of the main text).

An inocula bank from a donor simulation was used to initialize the individual's state. In order to mimic experimental practice, the individual's state after 48 h from the beginning of a donor simulation, starting from an initial cell concentration of $1x10^4$ cells/ml, was recorded and used later. Initial concentration of chemical species of the donor simulation was assumed to be the same as calculated above except for glucose, which was set to the glucose concentration of the Sabouraud broth (20 g/l).

B.4 Uptake-related parameters

From the maximum glucose uptake rate used by previous modelling efforts (Alexander and Jeffries, 1990; Steinmeyer and Shuler, 1989; Strässle et al., 1988), an initial range for the mean glucose uptake rate, u_{GLU}, was initially established at 13.3-20 glucose mmol/(g DW·h). This value proved to be insufficient to replicate the experimental glucose consumption curve, with levels of residual glucose in the medium remaining unrealistically high. For this reason, after adopting 20 mmol/(g DW·h) as the lower value of the interval, a high working value for the upper level, 120 mmol/(g DW·h) was established. This wide range allows the interval to include the most probable value of the parameter and, in turn, reflects the uncertainty in the parameter value. It is worth noting that in those works which refer to this value, the authors are

using different deterministic modelling approaches and, although the parameters are homologous, do not necessarily share the same parameter values. Additionally, this value is usually obtained from continuous culture experiments and thus is the mean uptake of the population growing in specific culture conditions, and not a hypothetic maximum individual uptake capacity, as u_{GLU} represents. Therefore, a population value, which is a consequence of all the individuals acting together, as well as the interactions between them and the abiotic components of the system, is being used to parameterize an inherently individual unaffected parameter. This drawback could be tackled by calibrating u_{GLU} using data from specifically designed glucose-limited continuous culture experiments. Discrepancies between population and individual-based approaches have been reported and explained with the Jensen's inequality (Hellweger and Bucci, 2009).

The mean uptake rate of organic N, u_{CN}, has been parameterized taking amino acids as a reference organic N compound. Steinmeyer and Shuler (1989) use three uptake systems with maximum uptake rates of 0.134, 0.182 and 0.281 g/(g DW·h) to represent the variety of both general and specific amino acid permeases present in *S. cerevisiae*. Assuming, respectively, the leucine, histidine, and lysine molar masses, the initial range of u_{CN} can be set to 1-4.5 mmol/(g DW·h). As opposed to glucose, this range was able to potentially reproduce experimental data. Although the same analysis of individual vs. population parameters also holds, experimental conditions (N limited cultures) and available data (no control of the uptaken N sources) could be hiding or mitigating this effect. Additionally, and taking into account the N limited-growth, initial temporal resolution of the data for the period of intense cell division (first two days approximately) will not allow a fine tuning of the model parameters related to the N uptake.

Sainz et al. (2003) use 300 mmol/(g DW·h) as the upper bound for the ammonium uptake rate. Since no more information was available, the initial range for the mean ammonium uptake rate, was heuristically set to 225-375 mmol/(g DW·h).

As previously stated, scar tissue is thought to be less efficient at facilitating the transport of molecules into the cell, thus the parameter ps represents a reduction on the yeast cell surface for nutrient exchange. Since the composition of the cell wall material at the centre of the chitin ring is unknown (Powell et al., 2003a), parameterization of p_s lies under the assumption that the reduction of surface area per cell scar ranges from the area occupied by the Chitin ring and the total area of the scar (chitin ring and central tissue), assuming a two-dimensional structure. Using data provided by Powell and co-workers (Powell et al., 2003a) (cell surface, mean bud scar diameter, mean bud scar surface, percentage of cell surface covered by bud scars, number of scars of the cell), and graphically estimating a chitin ring thickness of 0.35 μm, p_s initial range can be set to the interval 0.007-0.018 Cell scar^{-1}.

The parameter affecting glucose uptake due to the ethanol concentration in the medium, $p_{u\mathrm{OH}}$, can be considered as the resistance of the yeast to the toxic effects of the alcohol. This parameter can be approached from the maximum ethanol concentration a certain strain can bear based on previous experimental observations or knowledge. For *S. cerevisiae* and under oenological conditions, this level can be set around 15-16 % (vol/vol), although problematic stuck fermentations can appear from 13 % (vol/vol) (Ribéreau-Gayon et al., 1998). Assuming an ethanol density of 0.79 g/ml, this would lead to an alcohol concentration in the medium ranging from 102.7 to 126.4 g ethanol/l. However, the maximum ethanol concentration a strain can hold depends, although not exclusively, on the yeast strain, temperature, and aeration level (Ribéreau-Gayon et al., 1998). Thus, these reported levels are likely to be biased from a potential ethanol resistance represented by $p_{u\mathrm{OH}}$. Consequently, the initial range explored for the parameter has been set to 100-180 g ethanol/l.

The uptake and subsequent immobilization of unmetabolized organic compounds can change the C/N ratio of the yeast biomass. The model assumes that the uptake of C and N sources are switched on or off by this internal ratio in order to ensure that the individual's composition remains between typical values. Parameters governing this are r_{C} for the uptake of C sources, and r_{N} for the uptake of N sources. Based on previous INDISIM works (Gras et al., 2011), which assume a ratio of between 5 and 12 for a mixture of heterotrophic microorganisms, and since the yeasts have a higher relation of C to N sources than bacteria, then the assumed values are 6 for r_{C}, and 12 for r_{N}, which have been chosen and fixed as constants.

Since no previous information was available, standard deviation of the uptake-related parameters (σ_{GLU}, σ_{CN}, σ_{NH4}, σ_{ps}, $\sigma_{pu\mathrm{OH}}$) was obtained from the mean values of the parameters and a normalized standard deviation of 0.25 was assumed.

B.5 Metabolism-related parameters

Sainz et al. (2003) validate their model using wine fermentation data and use a maintenance energy of 1.4 mmol ATP/(g DW·h) for *S. cerevisiae* growing in an ethanol free medium. On the other hand, Jarzębski et al. (1989), after calibrating a continuous model to data obtained in a fermentation system, found maintenance energy requirements ranging from 0.21-0.27 glucose g/(g DW·h). Assuming that one mole of glucose yields 36 ATP or 2 ATP following, respectively, respirative and fermentative catabolic metabolisms, this gives a range for the parameter e of 0.12-0,27 glucose g/(g DW·h).

Sainz et al. (2003) take into account the effect of the concentration of ethanol on maintenance expenses of *S. cerevisiae* by means of a look up table partially based on experimental data. Subtracting maintenance expenses at y-intercept from the stated look up table showing maintenance expenses [mmol ATP/(g DW·h)] versus ethanol concentration (g/l), considering the maximum

and mean slope of the modified graph, and after appropriate transformation, a range for p_{OH} of 0.01 to 0.22 g glucose/ (g DW·h) can be set. Nevertheless, the upper value found was shown to be insufficient to reproduce experimental data, thus the upper level of the original range was set to a higher value to reflect its uncertainty. After this working hypothesis, the initial range for the parameter was set to 0.01-2 g glucose/(g DW·h). Since no previous information was available, standard deviation of the parameter, σ_{pOH}, was set from the mean value of the parameter and a normalized standard deviation of 0.25.

S. cerevisiae is a Crabtree-positive yeast and consequently even under aerobic conditions, fermentation predominates over respiration. This phenomenon has been observed in cultures growing over a glucose concentration of 9 g glucose/l (Ribéreau-Gayon et al., 1998). Since the parameter driving the accumulation of carbohydrate reserves, as external glucose is being exhausted by the yeast, l_S, must necessarily represent a lower glucose concentration than that observed experimentally in the Crabtree effect. Consequently, in the absence of more accurate data, l_S, has been estimated heuristically from the Crabtree glucose concentration and set to 0.05 to 4.5 g glucose/l.

In order to fully complete the metabolism of absorbed substrates, α_1 and α_2, the coefficients indicating the C moles required as energy source to synthesize a biomass unit following respectively stated Eqs. 2 and 3, have to be known. Using measurable experimental data, E_1, the metabolic energy required to create a biomass unit (mol of ATP/ mol CNMIC) from glucose and ammonium, can be estimated as follows:

$$E_1 = E_2 = \left(\frac{1-y}{y}\right) \cdot \frac{a_{yc}}{6} \cdot 2 \tag{B.1}$$

where y is the fermentative biomass yield production (g C-DW/g C-glucose). Recalling that the relation between these coefficients and the metabolic energy required to create a biomass unit is known (E_1=2/6 α_1), α_1 can be obtained. It is assumed that, E_2, the metabolic energy required to create a biomass unit from glucose and organic N is equal to E_1. Assuming a pure anaerobic growth of *S. cerevisiae*, y has been fixed to the value of 0.05 g DW/ g glucose (Alexander and Jeffries, 1990; Strässle et al., 1988).

According to the value given to the metabolism-related model parameters, Eq. A.9 and A.10 can be rewritten as Eq. B.2 and B.3:

$$150\ C_{GLU} + 1\ N_{NH4} \rightarrow 1\ C_6N + 96\ C_{OH} + 48\ C_{CO2} \tag{B.2}$$

$$145.5\ C_{GLU} + 4.5\ C_{CN} + 1\ N_{CN} \rightarrow 1\ C_6N + 96\ C_{OH} + 48\ C_{CO2} \tag{B.3}$$

which denotes the stoichiometry of coupled reactions needed to create a mole of simulated mass (C_6N) using, respectively, ammonium (NH_4) or organic nitrogen ($C_{4.5}N$).

B.6 Reproduction-related parameters

Johnston et al. (1977), studying coordination between growth and cell division of *S. cerevisiae*, found that small cells from a nitrogen-starved population of the strain X2180-1A must grow to a minimum volume of 25 μm^3 before bud emergence can occur, and that the execution of the early DNA-division cycle events for four *cdc* mutant strains is correlated with the attainment of a cell volume of approximately 30-35 μm^3. Johnston et al. (1979) determine cell volume at bud initiation of cells showing 0, 1, 2 or 3 bud scars and growing at different growth rates. From this data, a mean volume at bud initiation of virgin cells of 29.3 μm^3 with a mean normalized standard deviation of 0.16 can be computed. Finally, Takamatsu et al. (1985), after adjustment of simple population models in baker's yeast fed-batch cultures, found critical volumes at budding between 26-34.5 μm^3. Taking the range of 25-35 μm^3 as the minimum volume for a *S. cerevisiae* cell to start bud formation and assuming a cell density of about 1.1 g WW/ml (Bryan et al., 2010) and a moisture content of 75% (Briggs et al., 1981), the range explored for the m_C parameter is set to 6.87-9.63 pg DW. The standard deviation of the parameter, σ_{mC}, was set from the mean value of the parameter and a normalized standard deviation of 0.15.

According to the data of Johnston et al., (1979), obtaining the mean volume at bud initiation of cells with 2 scars and subtracting the mean volume at bud initiation of virgin cells, a mean volume increase of 6 μm^3 per cell cycle is obtained or 1.65 pg DW. Assuming that once in the budded phase, all the newly synthesized biomass goes to the daughter cell (Alberghina et al., 1998; Porro et al., 2009), this can be used as the central value of the model parameter Δm_{B1}. The final explored range for Δm_{B1} was set heuristically to 1.65±(0.41).

Strässle et al. (1988) assume that the minimum mass of a daughter cell is 6 pg DW which has been adopted as the central value for Δm_{B2}. As before, the range explored for the parameter was set to 6± (1.5). The standard deviation of the parameter, σ_{mB2}, was set from the mean value of the parameter and a normalized standard deviation of 0.25.

Johnston et al. (1979) worked with different glucose concentrations in a chemostat and found that slow growing haploid cells of *S. cerevisiae* showed a constant cell volume at division but that this volume increased progressively in cells growing over a specific growth rate of 0.17-0.23 h^{-1}. This is consistent with the assumption of a minimum time requirement and mass increase to release the bud, as assumed within the model. Thereby, a minimum mass increase

would limit slow growing cell division while a minimum time increase would limit fast growing cell division, allowing a progressively high volume increase. Therefore, budding time at the specific growth rate distinguishing between fast and slow growing cells (or the point where mass requirement equals time requirement) could be used as a value for the model parameter ΔT_2, the minimum time required to bud scission from bud initiation. Using empirically derived relationships between the length of the budded phase and doubling time (Hartwell and Unger, 1977; Tyson et al., 1979 and references therein) and the stated specific growth rates, the range of values studied for ΔT_2 has been set to 117-183 min. The standard deviation of the parameter ΔT_2, σ_{T2}, was set from the mean value of the parameter and a normalized standard deviation of 0.25.

B.7 Lifespan-related parameters

The stationary phase has been defined as the ability of yeasts to survive long periods (i.e. months) without added nutrients (Werner-Washburne et al., 1993). Keeping this in mind, and since the total cell count did not decrease during the length of the experiments, it has been assumed that cells do not die due to nutrient starvation but due to their limited replicative lifespan (see above). To reflect this assumption, the parameter t_d and its standard deviation, σ_{td}, have been fixed, respectively, at 1000 and 0 h.

An *S. cerevisiae* cell is only capable of undertaking a limited number of replications. The maximum replicative number of events a cell can experience, the Hayflick limit (Hayflick, 1968), is strain-specific. Once the Hayflick limit is reached, cells are incapable of further replication and enter into a state of senescence, which leads to cell death (Barker and Smart, 1996). Although it is strain dependent, the Hayflick limit for *S. cerevisiae* has been reported to be within the range of 9-30 cell divisions (Maskell et al., 2003 and references therein). Barker and Smart (1996) found a maximum value of 29 for a particular *S. cerevisiae* strain. Recalling that within a typical laboratory grown batch culture the expected fraction of seventh or greater generation cells is relatively low, < 1% (Powell et al., 2003b), the model parameter denoting maximum replicative lifespan, h_h, was fixed to the value of 30.

B.8 Parameters related to the actions over the medium

Substrates accounted for by the model affecting, either directly or indirectly, the fitness of the individuals are subjected to diffusive processes. Because of their different natures, the parameters driving these processes, d_j (units: fraction of translated particles), cannot be directly translated from the known, experimentally-measured diffusivities of the substrates (units: cm^2/s),

despite being related to them. Therefore, for purposes of the present contribution, the d_j values have been fixed assuming a glucose diffusivity in water at 25 °C of 6.9 x10^{-6} cm^2/s (Perry et al., 1999), organic N diffusivity in water at 25 °C of 8x10^{-6} cm^2/s (Longsworth, 1953), ammonium diffusivity in water at 30 °C of 18.6x10^{-6} cm^2/s (Kreft et al., 2001), ethanol diffusivity in water at 25 °C of 12.8x10^{-6} cm^2/s (Perry et al., 1999), and the value of d_{GLU} used within previous INDISIM works (Ginovart et al., 2005; Gras et al., 2010). From these values, the values adopted for d_j have been assumed to be 2.5x10^{-3} (j= GLU, CN) and 5.0x10^{-3} (j= NH4, OH).

References

Alberghina, L., Smeraldi, C., Ranzi, B.M, Porro, D., 1998. Control by nutrients of growth and cell cycle progression in budding yeast, analyzed by double-tag flow cytometry. *Journal of Bacteriology* 180, 3864-3872.

Alexander, M.A. and Jeffries, T.W., 1990. Respiratory efficiency and metabolite partitioning as regulatory phenomena in yeasts. *Enzyme and Microbial Technology* 12, 2-19.

Barker, M.G., Smart, K.A., 1996. Morphological changes associated with the cellular aging of a brewing yeast strain. *Journal of the American Society of Brewing Chemists* 54, 121-126.

Berry, D.R., 1988. Products of primary metabolic pathways. In *Physiology of Industrial Fungi* (ed. D.R. Berry), pp. 130-160. Blackwell Scientific Publications, Oxford.

Briggs D.E. et al., 1981. *Malting and brewing science* vol.1: Malt and sweet wort. 2nd ed., London ; New York : Chapman and Hall.

Bryan, A.K., Goranov, A., Amon, A., Manalis, S.R., 2010. Measurement of mass, density, and volume during the cell cycle of yeast. *Proceedings of the National Academy of Sciences of the U S A* 107, 999-1004. doi: 10.1073/pnas.0901851107.

Ginovart, M., López, D., Gras, A., 2005. Individual-based modelling of microbial activity to study mineralization of C and N and nitrification process in soil. *NonLinear Analalysis: Real World Applications* 6, 773-795. doi: 10.1016/j.nonrwa.2004.12.005

Gras, A., Ginovart, M., Portell, X., Baveye, P.C., 2010. Individual-based modelling of carbon and nitrogen dynamics in soils: parameterization and sensitivity analysis of abiotic components. *Soil Science* 175, 363-374. doi: 10.1097/SS.0b013e3181eda507

Gras, A., Ginovart, M., Valls, J., Baveye, J.C., 2011. Individual-based modelling of carbon and nitrogen dynamics in soils: parameterization and sensitivity analysis of microbial components. *Ecological Modelling* 222, 1998-2010. doi: 10.1016/j.ecolmodel.2011.03.009.

Hartwell, L. H., Unger, M.W., 1977. Unequal division in *Saccharomyces cerevisiae* and its implications for the control of cell division. *Journal of Cell Biology* 75, 422-435.

Hayflick, L., 1968. The limited in vitro lifetime of human diploid cell strains. *Experimental Cell Research* 37, 614-636. doi: 10.1016/0014-4827(65)90211-9

Hellweger, F.L., Bucci, V., 2009. A bunch of tiny individuals - Individual-based modeling for microbes. *Ecological Modelling* 220, 8-22. doi: 10.1016/j.ecolmodel.2008.09.004.

Jarzębski, A.B., Malinowski, J.J., Goma, G., 1989. Modeling of ethanol fermentation at high yeast concentrations. *Biotechnology and Bioengineering* 34, 1225-1230. doi: 10.1002/bit.260340915

Johnston, G.C., Pringle, J.R., Hartwell, L.H., 1977. Coordination of growth with cell division in the yeast *Saccharomyces cerevisiae*. *Experimental Cell Research* 105, 79-98. doi: 10.1016/0014-4827(77)90154-9

Johnston, G.C., Ehrhardt, C.W., Lorincz, A., Carter, B.L., 1979. Regulation of cell size in the yeast *Saccharomyces cerevisiae*. *Journal of Bacteriology* 137, 1-5.

Kreft J.U., Picioreanu, C., Wimpenny, J.W., van Loosdrecht, M.C., 2001. Individual-based modelling of biofilms. *Microbiology* 147, 2897-2912.

Longsworth L.G., 1953. Diffusion measurements, at 25 °C, of aqueous solutions of aminoacids, peptides and sugars. *Journal of the American Chemical Society* 75, 5705-5709. doi: 10.1021/ja01118a065

Maskell, D.L., Kennedy, A.I., Hodgson, J.A., Katherine, A.S., 2003. Chronological and replicative lifespan of polyploid *Saccharomyces cerevisiae* (syn. *S. pastorianus*). *FEMS Yeast Research* 3, 201-209. doi: 10.1016/S1567-1356(02)00199-X

Perry, R.G., Chilton, C.H., Kirkpatrick, S.D., 1999. *Chemical engineers handbook*. Chemical engineers handbook.

Powell, C.D., Quain, D.E., Smart, K.A., 2003a. Chitin scar breaks in aged *Saccharomyces cerevisiae*. Microbiology 149, 3129-3137. doi: 10.1099/mic.0.25940-0.

Powell, C.D., Quain, D.E., Smart, K.A., 2003b. The impact of brewing yeast cell age on fermentation performance, attenuation and flocculation. *FEMS Yeast Research* 3, 149-157. doi: 10.1016/S1567-1356(03)00002-3

Porro, D., Vai, M., Vanoni, M., Alberghina, L., Hatzis, C., 2009. Analysis and modeling of growing budding yeast populations at the single cell level. *Cytometry Part A* 75, 114-120. doi: 10.1002/cyto.a.20689.

Portell, X., Ginovart, M., Carbó, R., Gras, A., Vives-Rego, J., 2011. Population analysis of a commercial *Saccharomyces cerevisiae* wine yeast in batch culture by electric particle analysis, light diffraction and flow cytometry. *FEMS Yeast Research* 11, 18-28. doi: 10.1111/j.1567-1364.2010.00682.x

Ribéreau-Gayon, P., Dubourdieu, D., Donèche, B., 1998. *Traité d'oenologie* (6 Éd.): Tome 1. Microbiologie du vin. Vinifications. Paris: Dunod.

Sainz, J., Pizarro, F., Pérez-Correa, J.R., Agosin, E., 2003. Modeling of yeast metabolism and process dynamics in batch fermentation. *Biotechnology and Bioengineering* 81, 818-828. doi: 10.1002/bit.10535.

Scharlau, 2011. *Handbook of microbiological culture media*. 11th International edition, Barcelona, Spain.

Steinmeyer, D.E., Shuler, M.L., 1989. Structured model for *Saccharomyces cerevisiae*. *Chemical Engineering Science* 44, 2017-2030. doi: 10.1016/0009-2509(89)85138-3.

Strässle, C., Sonnleitner, B., Fiechter, A., 1988. A predictive model for the spontaneous synchronization of *Saccharomyces cerevisiae* grown in continuous culture. I. Concept. *Journal of Biotechnology* 7, 299-317. doi: 10.1016/0168-1656(88)90042-9.

Takamatsu, T., Shioya, S., Chikatani, H., Dairaku, K., 1985. Comparison of simple population models in a baker's yeast fed-batch culture. *Chemical Engineering Science* 40, 499-507. doi: 10.1016/0009-2509(85)85110-1

Tyson, C.B., Lord, P.G., Wheals, A.E., 1979. Dependency of size of *Saccharomyces cerevisiae* cells on growth rate. *Journal of Bacteriology* 138, 92-98.

Werner-Washburne, M., Braun, E., Johnston, G.C., Singer, R.A., 1993. Stationary phase in the yeast *Saccharomyces cerevisiae*. *Microbiological Reviews* 57, 383-401.

Appendix C

INDISIM-*Saccha* extension description

The model description follows the ODD (Overview, Design concepts, and Details) protocol for describing individual- and agent-based models (Grimm et al., 2006; 2010). To better follow this description, the reader should bear in mind the following: (i) individual and spatial cell state variables, which record the specific state of the system at each time step, are designated with a capital letter followed by a suffix; (ii) the model parameters are designated with a lower case letter; (iii) any particular random realizations following a given statistical distribution are designated with a lower case letter with a dieresis; and (iv) the modelled organic molecule are designated as CN with a suffix. Some of the important suffixes are: GLU (glucose), OH (ethanol), NH4 (ammonium), CO2 (carbon dioxide), O2 (molecular oxygen) and MIC (from a microorganism biomass). SC will stand for Spatial Cell

C.1 Purpose

The model has been developed to analyse the dynamics of *S. cerevisiae* batch cultures evolving into a stirred aerobic liquid medium with glucose as a main C source, organic and inorganic N sources, and with culture conditions not promoting a noticeable Crabtree effect (i.e., low initial glucose).

C.2 Entities, state variables, and scales

The model takes into account three entities: *S. cerevisiae* individual cells, spatial cells, and the environment. A yeast cell (I_i) is defined by the variables: $X(t)$, $Y(t)$, and $Z(t)$, identifying its position in the domain; $M(t)$, its structural mass (CN_{MIC}-pmol); $B(t)$, its genealogical age (bud scars); $P(t)$, the reproduction phase in the cellular cycle in which the cell is currently (unbudded or budding phase); $M_{Start}(t)$, its "Start mass" (CN_{MIC}-pmol), the mass required to change from the unbudded to the budding phase; $M_{inc}(t)$, the increased mass (CN_{MIC}-pmol) since the cell entered to the budding phase; $T_{inc}(t)$, time spent into the current reproduction phase (time steps); $R_{GLU}(t)$, the amount of C stored in the cell as reserve carbohydrates or in the model as a glucose polymers (glucose-pmol), $R_{CN}(t)$, the amount of organic N stored in the cell as a reserve (CN-pmol); $C^{in}_{GLU}(t)$, amount of non-metabolized glucose inside the cell (glucose-pmol); $D(t)$ the mortality index to evaluate cell viability; and $T_a(t)$ the individual adaptation time to new environmental conditions (individual lag time). Letting $I=I(t)$ denote the number of individuals at time t, and identifying an individual by i, the population's state at t is:

$$I_n = \left\{ I_i \left[X, Y, Z, M, B, P, M_{Start}, M_{inc}, T_{inc}, R_{GLU}, R_{CN}, C^{in}_{GLU}, D, T_a \right] \right\}_{i=1,2,\ldots,n} \quad \text{(C.1)}$$

The simulated space ($Grid_{3D}$) is a cube which is divided into spatial cells (S_{xyz}) of equal size and shape described by a vector defined by the variables: $S_{GLU}(t)$ the amount of glucose (pmol); $S_{CN}(t)$, the amount of organic N (pmol); $S_{NH4}(t)$, the amount of ammonium (pmol); $S_{OH}(t)$, the amount of ethanol

(pmol); $S_{O2}(t)$, the amount of molecular oxygen; and $S_{CO2}(t)$, the amount of carbon dioxide (pmol). The whole three dimensional grid is then described by:

$$Grid_{3D} = \{S_{xyz}[S_{GLU}, S_{CN}, S_{NH4}, S_{OH}, S_{O2}, S_{CO2}]\}_{x=1,\ldots,iq;y=1,\ldots,iq;z=1,\ldots,iq} \quad \text{(C.2)}$$

The environment simulates a stirred liquid medium closed in a cube whose faces do not allow the ingress or egress of either organic or inorganic elements, with the exception of molecular oxygen, which may enter to the medium. The stirring process causes a random change in the position of the individuals, and a redistribution of the inorganic elements.

The temporal evolution of the system is divided into equal intervals associated with time steps (TS).

C.3 Process overview and scheduling

Global simulation scheduling is mainly made up of four sections: (i) the initialization of the simulated system, when the input data is entered, (ii) the setting of the initial configuration of the population, (iii) the initial setting of the space, and (iv) the time step loop which is repeated until the end of the defined time steps. The time step loop scheduling includes, chronologically: (iv.i) the random order of the individuals' acting order, which will act one behind the other, (iv.ii) the individual actions loop (sections C.7.1-C.7.4), (iv.iii) the actions on the medium (section C.7.5), and (iv.iv) output of the desired aggregated and state variables.

At each time step and at the individual actions loop, existing yeast cells perform, sequentially, the following set of actions: uptake (section C.7.1), metabolism (section C.7.2), reproduction (section C.7.3), and lifespan (section C.7.4). Yeast activity is locally described; therefore actions of individuals take place and only affect the spatial cell they are in. Variables are updated as soon as their value is calculated by a process (asynchronous updating).

C.4 Design concepts

Emergence. The main phenomena to emerge from the model are the evolution of the nutrient concentrations over time, the population growth behaviour, and the distribution of individual properties at particular times through the system evolution, emphasizing mass and scar distributions as a direct simulation homologous to well-established experimental techniques. The phenomena described above are not imposed by model rules but are consequences of individual behaviour and interactions between biotic and abiotic elements.

Adaptation. The model assumes that the individual yeast changes its metabolism as an adaptation strategy to the glucose depletion. The individual begins to store carbohydrates as C reserve that will be required to survive during starvation conditions (Walker, 1998).

Sensing. Individual yeasts are assumed to sense medium glucose concentration and change their internal metabolic fluxes to adapt to the changing environmental conditions. This sensing mechanism is not explicitly modelled but individuals are simply assumed to know external glucose.

Interaction. Both intraspecific direct and indirect interaction phenomena are taking place within the modelled system. Indirect interactions include competition for nutrient sources (glucose, organic N, ammonium, and oxygen). Detrimental effects into individual fitness due to external ethanol content, which is a main product of the individual yeast metabolism, can be seen as a direct interaction among yeast individuals.

Stochasticity. Stochasticity is introduced into the model when setting some parameter values of the individuals using a positive truncated normal distribution. Randomness is also considered when the rules are applied to individuals and to spatial cells by using probabilistic distributions to deal with or manage individual events. Random processes or events include: (i) determination of the individual lag phase; (ii) maximum uptake rate at a given time interval; (iii) computation of the toxic effect of the ethanol on the glucose uptake; (iv) computation of maintenance energy requirements; (v) yeast start mass assignation; (vi) checking successful completion of budded phase (both within expended time and accumulated cell mass); (vii) testing for cell death; and (viii) random redistribution of the individual position. The sequence of action of individuals changes randomly at each time step in order to avoid privileging first-acting cells.

Observation. The behaviour of the system is followed by means of (i) data related to the global properties of the system and (ii) data concerning properties of individual yeast cells. The former may include, but are not limited to, information about the temporal evolution of the amount of nutrients (glucose, organic N, and ammonium), the amount of metabolites (ethanol and carbon dioxide), the average nutrient consumption, the number of viable yeast cells, the number of dead yeast cells, viable yeast biomass, dead yeast biomass, heat dissipation of the system, maintenance energy expended by the yeast population (defined as the number of metabolized nutrient particles not used in the production of new biomass), and the mean mass of the cell population. In addition, it is feasible to obtain the evolution of microscopic population parameters, which are controlled at individual level, such as, among other potential individual properties, the distributions of genealogical age and of the mass of the individuals. These distributions are related mainly to the cellular cycles of budding reproduction, and reflect the state of the yeast population at given times in the process. Because the simulator saves information about every cell at each time step, this recorded information makes it possible to construct bar charts and histograms to represent the latter distributions. These simulation outputs show the time evolution of the structure of the population throughout the yeast culture. The preceding separation of output simulation results mirrors the classification of experimental techniques used to study said properties.

C.5 Initialization

Following the same methodology as Gras and co-workers (Ginovart et al., 2005; Gras et al., 2011), individual yeast mass is expressed in molar units, so that the yeast biomass elemental formula may be expressed as $C_{a_{yC}}N$. Additionally, to make the simulator easy for unskilled users to operate, the application of experimental units into the inputter section is desirable and, consequently, they have to be converted to simulation units (molar).

The initial number of individuals of a given simulation, $I(t{=}0)$, is calculated according to a provided initial cell density (individuals/ml) and the simulation volume. n_0 individuals are initialized as summarized in Table C.1.

Table C.1. Initialization of individual state variables at the beginning of the simulation used as inoculum. Within the table, a denotes the initial value, iq is the number of spatial cells per dimension, m_C is the critical mass at reproduction model parameter, and t_a is the mean individual lag time parameter. Standard deviations of the model parameters are denoted by σ.

Individual state variable	Initial value
M	$a = m_C \cdot U(0.8,1.2)$
M_{Start}	$a; a \rightarrow N(m_C, \sigma_{mc})$
B	0
P	1
M_{inc}, T_{inc}, R_{GLU}, R_{CN}, $\boldsymbol{C_{GLU}^{in}}$, D	0
T_a	$Max(0, a); a \rightarrow N(t_a, \sigma_{ta})$
X,Y,Z	$a\,; P[a = k] = \frac{1}{iq}; k = 1, \ldots, iq$

Spatial cells state variables are initialized according to an initial concentration (g/l) which is equally distributed among all spatial cells. Substrate molar mass, simulated volume and the number of spatial cells allow changing from initial concentration to initial cell content (pmol per spatial cell). Organic labile N is assumed to have the following elemental formula: $C_{a_C}N$.

C.6 Input data

The model does not use input data to represent time-varying processes.

C.7 Submodels

C.7.1 Uptake submodel

Four substrates are taken into account by the model: glucose, organic N, ammonium, and molecular oxygen (*j*=GLU, CN, NH4, O2). As previous INDISIM models (Ginovart et al., 2005; Gras et al., 2011), it is assumed that the maximum individual uptake of C and N sources is controlled by the internal C to N ratio (Eq. C.3), a value lower than or equal to r_C to uptake C sources and greater than or equal to r_N to uptake N sources.

$$C/_N(t) = \frac{M(t)\cdot a_{yc} + R_{GLU}(t)\cdot 6 + R_{CN}(t)\cdot a_C + C_{GLU}(t)\cdot 6}{M(t) + R_{CN}} \quad \text{(C.3)}$$

where a_C is the C to N molar ratio of the simulated organic labile compound (CN).

The maximum individual uptake, $U^j_{MAX}(t)$, in a time step for the substrate *j* (pmol), will be determined by the size of the individual (i.e. its structural mass). As scar tissue is thought to be less efficient than normal cell wall material at facilitating the transport of macromolecules into the cell (Powell et al., 2003a), it has been previously proposed that an increase in the bud scars number can limit the surface area for nutrient exchange (Mortimer and Johnston, 1959). Interestingly, as shown by Powell et al. (2003a), bud scars (as birth scar) expand with genealogical age and, as a result, the percentage of cell surface occupied by a scar is almost constant throughout the genealogical lifespan of the cell, making the effects of the number of scars of the cell, $B(t)$, constant (Eqs. C.4 .and C.5). It is worth noting that this may not be true if the individual's mass, and not its surface area, is being considered in its own right. Assuming a characteristic cell shape and density, the relationship between the cell surface area, cell volume, and cell mass, which is usually used to obtain experimentally the unitary uptake rate, can be approached (see, for instance, Ginovart et al., 2005; Gras et al., 2010, 2011). Keeping in mind the high level of uncertainty of the necessary estimates (which makes this computation highly uncertain), this unit transformation has been avoided. Nonetheless, a constant effect of the number of scars on the uptake has been taken into account. Additionally, in the case of glucose (Eq. C.5), the model assumes that the maximum uptake is also affected by the extracellular ethanol content of the spatial cell the yeast is currently in. This toxic effect of the ethanol has already been taken into account in previous modelling efforts (Steinmeyer and Shuler, 1989; Sainz et al., 2003).

$$U^j_{MAX}(t) = \ddot{u}_j \cdot M(t) \cdot \big(1 - p_s\, B(t)\big); j = \mathrm{CN}, \mathrm{NH4}, \mathrm{O2} \quad \text{(C.4)}$$

$$U^j_{MAX}(t) = \ddot{u}_j \cdot M(t) \cdot \big(1 - p_s\, B(t)\big) \cdot \left(1 - \frac{S_{OH}(t)}{\ddot{p}_{uOH}}\right); j = \mathrm{GLU} \quad \text{(C.5)}$$

where $\ddot{u}_j$ is the specific substrate j uptake rate (pmol/pmol CN_{MIC}/TS), p_s is a penalization per cell scar and $\ddot{p}_{uOH}$ is the tolerance to ethanol of the yeast (pmol Ethanol/SC). $\ddot{u}_j$ is a non-negative random draw from a normal distribution with mean u_j and standard deviation σ_{uj}. $\ddot{p}_{uOH}$ is a non-negative random draw from a normal distribution with mean p_{uOH} and standard deviation σ_{puOH}. Negative numbers for $U^j_{MAX}(t)$ are not allowed and the value of zero is used instead.

The substrate uptake is also influenced by the amount of available substrate, that is the amount of substrate j present in the spatial cell, $S_j(t)$, in which the yeast is evolving. The substrate j uptake at a given time step, $U_j(t)$, is then the minimum quantity between the $U^j_{MAX}(t)$ and $S_j(t)$.

C.7.2 Metabolism submodel

For clarity, the metabolism submodel is divided into these sections: (i) variables update, (ii) maintenance requirements, (iii) carbon reserves creation section, (iv) new mass synthesis, (v) substance release, and (vi) individual lag phase modelling.

(i) *Variables update*. The internal amount of glucose able to be metabolized is given by its glucose uptake in a time step and the amount of non-metabolized glucose remaining from the previous time step, therefore, the glucose uptake is modified as follows: $U_{GLU}(t) = U_{GLU}(t) + C^{in}_{GLU}(t-1)$. Similarly, organic N uptake is also recalculated to take into account cell's internal reserves: $U_{CN}(t) = U_{CN}(t) + R_{CN}(t\text{-}1)$.

(ii) Maintenance requirements.

The model assumes the respirative catabolic pathway (glycolysis and Krebs cycle) as the first option in achieving metabolic energy. Nevertheless, since respiration is totally optional for *S. cerevisiae* (Alexander and Jeffries, 1990), it is assumed that the cell is also capable of using the fermentative catabolic pathway (glycolysis and alcoholic fermentation) once the uptaken O_2 is depleted. The following widely-accepted (Madigan, 2006) coupled global reaction for the respirative (Eq. C.6) and fermentative (Eq. C.7) catabolic pathways allows us to establish the metabolism model:

$$C_6H_{12}O_6 + 6\,O_2 \rightarrow 6\,H_2O + 6\,CO_2 + 36\,ATP \quad \text{(C.6)}$$

$$C_6H_{12}O_6 \rightarrow 2\,C_2H_6O + 2\,CO_2 + 2\,ATP \quad \text{(C.7)}$$

A yeast cell requires energy for its cellular maintenance, and it is assumed that this energy is proportional to its structural mass, $M(t)$. Additionally, the ethanol accumulated during the fermentation is a potent chemical stress factor for yeast cells so various physiological adaptations, which are thought to protect it against ethanol, occur in the cell (Walker, 1998), and measured maintenance requirements are known to be increased (Yang et al., 2012). Hence, if e_R is the maintenance rate through respirative metabolism (pmol

glucose/pmol of CN_{MIC}/TS) that a cell needs to remain viable for the current time step, we assume that the total energy required to maintain viability by respiring glucose, $Ma(t)$, is given by Eq. C.8:

$$Ma(t) = M(t) \cdot \left(e_R + \ddot{p}_{\mathrm{OH}}\, S_{\mathrm{OH}}(t)\right) \quad \text{(C.8)}$$

where $\ddot{p}_{eOH}$ is a penalization due to local ethanol content (pmol of glucose/pmol of CN_{MIC}/(pmol of ethanol/SC)). $\ddot{p}_{OH}$ is a non-negative random draw from a normal distribution with mean p_{OH} and standard deviation $\sigma_{p\mathrm{OH}}$. The model assumes that maintenance energy must be covered first before a cell is able to experience growth, i.e. new mass synthesis.

Taking Eq. C.8 into account, if $U_{\mathrm{GLU}}(t) \geq Ma(t)$, so the cell has enough glucose to cover Ma, it remains viable, and $U_{\mathrm{GLU}}(t)$ is updated by subtracting Ma. The remaining $U_{\mathrm{GLU}}(t)$ will then be used to create C reserves or new mass (see below). In case $U_{\mathrm{GLU}}(t) < Ma(t)$, then the cell does not have enough glucose to cover Ma and it needs to use its C reserves, $R_{\mathrm{GLU}}(t)$. In this case, it is assumed that a cell is not able to create new biomass. Individuals that are not able to fulfil Ma, may die (see section C.7.4. lifespan submodel below). Individuals that have completed their maintenance and still have glucose inside the cell are able to use it in the creation of new mass or new C reserves.

Notice that respiration of a glucose mol requires oxygen (Eq. C.7). The model assumes that if $U_{\mathrm{O2}}(t)$, the uptaken O_2 in the current time step, is not enough to fulfil Eq. C.7 stoichiometry, fermentative catabolism is followed. In this case, the amount of non-covered maintenance energy has to be achieved under fermentative conditions using a ratio, $\epsilon_{R/F}$ (mol fermented/mol respired), which is the relationship between the energy obtained from both catabolic pathways.

(iii) *Carbon reserve creation*. Resting and stressed cells accumulate high concentrations of storage carbohydrates. The main storage carbohydrate in yeast cells are trehalose and glycogen (Panek, 1991). Glycogen degradation and synthesis are controlled by environmental factors (e.g. Rowen et al., 1992). Once exogenous sugars and other essential nutrients become limiting, yeast can accumulate glycogen when their cell cycle is arrested. Using the stored glycogen under starvation conditions by endogenous fermentation contributes to cellular maintenance by providing ATP to keep cell viability (Walker, 1998). During the initial stages of fermentation of brewer's wort, *S. cerevisiae* rapidly degrades stored glycogen. To model this behaviour, it is assumed that reserve carbohydrates are created, and thus $R_{\mathrm{GLU}}(t+1) = U_{\mathrm{GLU}}(t)$, when glucose content in the spatial cell is below a threshold (I_S). Otherwise, new mass can be synthesized.

(iv) *New mass synthesis*. In order to synthesize new mass, *S. cerevisiae* needs a C source, an N source and a certain amount of energy. Yeast anabolic reactions (energy-consuming, reductive process) have to be coupled to catabolic reactions (energy-giving, oxidative process), which is achieved by dehydrogenase enzymes which predominantly use NADP and NAD,

respectively, as redox cofactors (Walker, 1998). Combing stated respirative catabolic (Eq. C.6) with the anabolic needs of C, N, O_2, and energy, Eq. C.9 and C.10 can be obtained.

$$a_{yc}\,\mathrm{C_{GLU}} + 1\,\mathrm{N_{NH4}} + E_{R1} \rightarrow a_{yC}\,\mathrm{C}_{CN_{MIc}} + 1\,N_{CN_{MIc}}$$
$$\alpha_{R1}\,\mathrm{C_{GLU}} + \alpha_{R1}\,\mathrm{O_{GLU}} + \alpha_{R1}2\,O_{O2} \rightarrow \alpha_{R1}\,\mathrm{O_{H2O}} + \alpha_{R1}\,\mathrm{C_{CO2}} + \alpha_{R1}2\,\mathrm{O_{CO2}} + \frac{36}{6}\,\alpha_{R1}$$
$$E_{R1} = \frac{36}{6}\,\alpha_{R1}$$

$$\left(a_{yc} + \alpha_{R1}\right)\mathrm{C_{GLU}} + 1\,\mathrm{N_{NH4}} + \alpha_{R1}\,\mathrm{O_{GLU}} + \alpha_{R1}2\,O_{O2} \rightarrow a_{yC}\,\mathrm{C}_{CN_{MIc}} + 1\,N_{CN_{MIc}} + \alpha_{R1}\,\mathrm{O_{H2O}} + \alpha_{R1}\,\mathrm{C_{CO2}} + 2\,\alpha_{R1}\,\mathrm{O_{CO2}} \quad \text{(C.9)}$$

$$\left(a_{yc} - a_c\right)\mathrm{C_{GLU}} + a_c\,\mathrm{C_{CN}} + 1\,\mathrm{N_{CN}} + E_{R2} \rightarrow a_{yC}\,\mathrm{C}_{CN_{MIc}} + 1\,N_{CN_{MIc}}$$
$$\alpha_{R2}\,\mathrm{C_{GLU}} + \alpha_{R2}\,\mathrm{O_{GLU}} + \alpha_{R2}2\,O_{O2} \rightarrow \alpha_{R2}\,\mathrm{O_{H2O}} + \alpha_{R2}\,\mathrm{C_{CO2}} + \alpha_{R2}2\,\mathrm{O_{CO2}} + 6\,\alpha_{R2}$$
$$E_{R2} = 6\,\alpha_{R2}$$

$$\left(a_{yc} - a_c + \alpha_{R2}\right)\mathrm{C_{GLU}} + a_c\,\mathrm{C_{CN}} + 1\,\mathrm{N_{CN}} + \alpha_{R2}\,\mathrm{O_{GLU}} + \alpha_{R2}2\,O_{O2} \rightarrow a_{yC}\,\mathrm{C}_{CN_{MIc}} + 1\,N_{CN_{MIc}} + \alpha_{R2}\,\mathrm{O_{H2O}} + \alpha_{R2}\,\mathrm{C_{CO2}} + \alpha_{R2}2\,\mathrm{O_{CO2}} \quad \text{(C.10)}$$

Similarly, combing stated fermentative catabolic (Eq. C.7) with the anabolic needs of C, N, and energy, Eq. C.11 and C.12 can be obtained

$$a_{yc}\,\mathrm{C_{GLU}} + 1\,\mathrm{N_{NH4}} + E_{F1} \rightarrow a_{yC}\,\mathrm{C}_{CN_{MIc}} + 1\,N_{CN_{MIc}}$$
$$\alpha_{F1}\,\mathrm{C_{GLU}} \rightarrow \alpha_{F1}\,\frac{4}{6}\,\mathrm{C_{OH}} + \alpha_{F1}\,\frac{2}{6}\,\mathrm{C_{CO2}} + \frac{2}{6}\,\alpha_{F1}$$
$$E_{F1} = \frac{2}{6}\,\alpha_{F1}$$

$$\left(a_{yc} + \alpha_{F1}\right)\mathrm{C_{GLU}} + 1\,\mathrm{N_{NH4}} \rightarrow 1\,\mathrm{C}_{a_{yc}}\mathrm{N} + \frac{4}{6}\alpha_{F1}\,\mathrm{C_{OH}} + \frac{2\alpha_{F1}}{6}\mathrm{C_{CO2}} \quad \text{(C.11)}$$

$$\left(a_{yc} - a_c\right)\mathrm{C_{GLU}} + a_c\,\mathrm{C_{CN}} + 1\,\mathrm{N_{CN}} + E_{F2} \rightarrow a_{yC}\,\mathrm{C}_{CN_{MIc}} + 1\,N_{CN_{MIc}}$$
$$\alpha_{F2}\,\mathrm{C_{GLU}} \rightarrow \alpha_{F2}\,\frac{4}{6}\,\mathrm{C_{OH}} + \alpha_{F2}\,\frac{2}{6}\,\mathrm{C_{CO2}} + \frac{2}{6}\,\alpha_{F2}$$
$$E_{F2} = \frac{2}{6}\,\alpha_{F2}$$

$$\left(a_{yc} - a_c + \alpha_{F2}\right)\mathrm{C_{GLU}} + a_c\,\mathrm{C_{CN}} + 1\,\mathrm{N_{CN}} \rightarrow a_{yC}\,\mathrm{C}_{CN_{MIc}} + 1\,N_{CN_{MIc}} + \frac{4}{6}\alpha_{F2}\,\mathrm{C_{OH}} + \frac{2\alpha_{F2}}{6}\mathrm{C_{CO2}} \quad \text{(C.12)}$$

where in Eqs. C.9-C.12 $\mathrm{C_{GLU}}$ are C moles from glucose; $\mathrm{C_{CN}}$ are C moles from the modelled organic N; $\mathrm{N_{NH4}}$ are N moles from ammonium; $\mathrm{N_{CN}}$ are N moles from the modelled organic N; $\mathrm{C}_{CN_{MIc}}$ are C moles of yeast biomass; $\mathrm{N}_{CN_{MIc}}$ are N moles of yeast biomass; $\mathrm{C_{OH}}$ are C moles from the ethanol; $\mathrm{C_{CO2}}$ are C moles from the carbon dioxide; $\mathrm{O_{H2O}}$ are O moles from the water; $\mathrm{O_{O2}}$ are O moles from the molecular oxygen; α (mol C-glucose/mol $\mathrm{CN_{MIC}}$) is a parameter related to the yield and indicates the C moles from glucose as the energy source needed to synthetize a mass unit, whose value is different depending on the metabolic pathway followed: α_{R1} and α_{R2} following respirative metabolism, and α_{F1} and α_{F2} following fermentative metabolism; E (mol ATP) is metabolic

energy required to create a mass unit, E_{R1} and E_{R2} are defined when respirative metabolism is followed, and E_{F1} and E_{F2} under fermentative metabolism.

Taking into account that energy is usually stored inside microorganisms as macromolecules so the amount of ATP remains low (Madigan, 2006), Eqs. C.9-C.12 assume that the addition of metabolic energy from anabolic and catabolic subreactions becomes zero.

Knowing the remaining amounts of C inside the cell, $U_{GLU}(t)$; organic N, $U_{CN}(t)$, ammonium $U_{NH4}(t)$, and $U_{O2}(t)$, the C and N composition of the substrates and the stoichiometry of the global reactions, the moles of new mass, Δm, created will be defined by the minimum value found multiplying the amount of reactants by the C (or N or O2) moles of the substrate divided by the corresponding stoichiometric coefficient. Finally, the submodel assumes that a cell will synthesize new mass consecutively following Eqs. C.9-C.12 . Once the newly created mass is known, the individual mass is set to: $M(t)=M(t\text{-}1)+\Delta m$. Similarly, if the individual is budding, the increased mass since the beginning of the budding phase, M_{inc}, is also updated $M_{inc}(t)=M_{inc}\ (t\text{-}1)+\Delta m$. If N sources limit the mass synthesis, and thus some glucose still remains inside the cell, it is stored in the corresponding individual state variable. Therefore, $C^{in}_{GLU}(t) = U_{GLU}(t)$.

(v) *Substance release*. As a consequence of mass creation, the amount of fermented glucose is also known. From this, metabolic heat and products are released to the spatial cell. The non-used ammonium and molecular oxygen are released to the medium but, taking into account that yeast cells store large pools of endogenous amino acids (Eddy, 1980; Steinmeyer and Shuler, 1989), the model considers that remaining organic N is stored as internal N reserves, R_{CN}. Therefore, $R_{CN}(t)= U_{CN}(t)$.

(vi) *Modelling the individual lag phase*. The population lag phase is a period of zero growth that appears when inoculated cells experience a change in nutritional status, alterations in physical growth conditions, or when cells need to recover from damages typical of stressing conditions (e.g., a previous stationary phase). This last case has been assimilated into the model and modelled in a simple way. Inoculum cells are assumed to have an individual lag time, T_a (in TS), which has to be overcame before starting cell growth (i.e., biomass synthesis) or, in practice, following the metabolism described in points i to v above. Until T_a is not surpassed, the cells only are able to cover maintenance as described above by using uptaken substrates. No accumulation of uptaken substrates is permitted.

C.7.3 Reproduction submodel

The model assumes that a cell can start a new cell cycle only after the preceding cycle is completed, and that the cellular cycle involves two differentiated phases. Phase 1, or unbudded phase ($P(t)=1$), covers most of phase G1 and a very small fraction of phase S in the traditional division of the

cell cycle; while phase 2, or budding phase ($P(t)=2$), covers a small fraction of G1, most of S and all of G2 and M (see Ginovart and Cañadas, 2008; Prats et al., 2010; Ginovart et al. 2011a;b, and references therein).

Conceptually, the model assumes that in the unbudded phase the yeast cell is getting ready for budding and that change to the budding phase takes place only when the cell: (i) has attained a minimum cellular mass, defined by the parameter m_C, the critical or minimum mass; and (ii) has achieved a minimum growth of its mass, which is related to the model parameter Δm_{B1}. Note that within the model, Phase 1 does not need to be completed in a given time interval. In practice, this behaviour is achieved by comparing, every time step, if the cell mass $M(t)$, has attained its needed mass at "Start", $M_{Start}(t)$, which is set at the beginning of Phase 1 so it fulfils the stated conceptual assumptions (see computation of M_{Start} below). Therefore, if, at a given time step, $M(t) \geq M_{Start}(t)$, the cell enters the budding phase. Note that this check in the model is to see whether an individual cell has reached a "Start" mass, irrespective of its original value and growth rate.

The budding phase is the least flexible in the cellular cycle as it requires both temporal and growth checks. Within the model, two conditions must be satisfied for the releasing of the bud, and the subsequent change to the unbudded phase. These are: (i) a minimum growth of mass, which is related to the parameter Δm_{B2}; and (ii) a minimum time interval, which is related to the parameter Δt_2. The first requirement is met when $M_{inc} \geq \ddot{\Delta m}_{B2}$; and the second when $T_{inc} > \ddot{\Delta t}_{B2}$, where $\ddot{\Delta m}_{B2}$ and $\ddot{\Delta t}_{B2}$ are non-negative random draws following, respectively, $N(\Delta m_{B2}, \sigma_{B2})$ and $N(\Delta t_2, \sigma_{t2})$. The first condition is necessary because a yeast cell must have a minimum number of molecules and achieve minimum structural mass in order to function as an independent entity. On the other hand, the bud growth, even under optimal growth conditions, needs a minimum time interval to be completed; this is represented by the second condition.

The budding phase is completed with the cell division. The mass of the mother cell after the cell division, $M(t+1)$, is equal to the actual mass less the daughter mass (m_D) as follows: $M(t+1)=M(t)-m_D$. As previously assumed in other modelling efforts (Porro et al., 2009), once the cell passes "Start", all the newly synthesized mass goes to the daughter cell, which has been experimentally corroborated (Alberghina et al., 1998 and references therein). Thus, $m_D=M_{inc}(t)$, represents the accumulated mass since the start of the budding phase until t.

Once m_D and $M(t+1)$ are known, and since both mother and daughter cells start a new unbudded phase, their mass at "Start", $M_{Start}(t+1)$, have to be computed. Taking the mother cell as an example, If $M(t+1) \leq m_C - \Delta m_{B1}$ then the "Start" mass assigned to the cell is $M_{Start}(t+1)= \ddot{m}_C$; otherwise, if $M(t+1) > m_C - \Delta m_{B1}$ then the "Start" mass assigned to the cell is $M_{Start}(t+$

$1) = M(t+1) + \Delta\ddot{m}_{B1}$where $\ddot{m}_C$ and $\Delta\ddot{m}_{B1}$ are positive random draws following, respectively, N(m_C, σ_{mC}) and N(Δm_{B1}, σ_{mC}). The calculus for the daughter cell is analogous. Notice that using this method, the previously stated conceptual assumptions of the requirements of the unbudded phase are met.

Properties of the, until now, mother cell are shared with the new daughter cell as a result of the bud detachment. The model assumes that this distribution of properties is proportional to the fraction of the detached mass. Therefore, and as previously stated, the fraction of the detached mass can be obtained from the relation:

$$\frac{M(t+1)}{M(t)} + \frac{m_D}{M(t)} = 1 \tag{C.13}$$

Using the previous expression and knowing $R_{\mathrm{GLU}}(t)$, $R_{CN}(t)$ and $C^{in}_{GLU}(t)$ for the mother cell, $R_{\mathrm{GLU}}(t+1)$, $R_{\mathrm{CN}}(t+1)$ and $C^{in}_{\mathrm{GLU}}(t+1)$ for both the mother and the daughter cell can be easily computed.

The daughter cell (a virgin cell since its number of bud scars is 0) is born in the same spatial cell occupied by the mother. Moreover, since the previous history of the mother cell is affecting both the cell and the bud, $D(t+1)$, the time without covering maintenance requirements of the daughter cell is also inherited from the main cell at detachment. The daughter cell remains active since the reproduction process has finished but it does not act until the next time step is reached.

C.7.4 Lifespan submodel

According to Maskell et al. (2003), it is essential to define the chronological lifespan as a consequence of cumulative and irreversible damage to intracellular components during extended stationary phase, which compromises cell integrity and leads to death and autolysis. In contrast, replicative lifespan is related to the number of divisions that an individual cell undertakes before entering a non-replicative state termed senescence, which leads to cell death and autolysis (Maskell et al., 2003; Jenkins et al., 2003 and references therein).

Both chronological and replicative lifespan are taken into account within the model. Chronological lifespan is followed using a mortality index, $D(t)$, which satisfies the following rules: (i) whenever a cell is not able to satisfy its maintenance requirements in a time step, the simulator increases the mortality index of the cell by one: $\mathrm{D}(t+1) = D(t)+1$; conversely, (ii) the index is set to 0, $D(t+1)=0$. Chronological lifespan is evaluated at every time step by comparing $D(t)$ and $\ddot{t}_d$, a non-negative random draw from an N(t_d, σ_{td}). Here, $\mathrm{t_d}$ denotes an average time beyond which the cell cannot survive. Thus, if $\mathrm{D}(t)>t_d$, the cell dies. In the model the individual cells may die, indirectly, due to ethanol excess or low glucose concentration. These reduce the ability of the yeast to cover its maintenance energy and increase mass. The magnitude of these unfavourable conditions will determine the vitality and viability of the individual yeast cell.

Replicative lifespan is followed using the cell's replicative age (i.e. the number of cell scars), $B(t)$. Once a cell reaches the maximum number of scars it can experience, known as the Hayflick limit (Hayflick, 1968; Jenkins et al. 2003, and references therein), h_l, the yeast cell dies. When the cell dies, its mass remains in the medium and no turnover is modelled.

C.7.5 Actions on the medium

Action on the medium includes simulation of the stirring process and of the action of sodium thioglycolate, which is commonly used as a reducing agent for molecular oxygen in laboratory practices.

The stirring process implies an equiprobable random change of an individual's position within the spatial domain; input of external oxygen (see below); and the uniform redistribution of the existing substrates, once all the individuals have acted, as follows:

$$\left\{ s_{(j)}^{xyz}(t+1) = \frac{\sum s_{(j)}^{xyz}(t)}{iq^3} \right\}_{x=1,\dots,iq;y=1,\dots,iq;z=1,\dots,iq} \tag{C.14}$$

where, $s_{(j)}^{xyz}$ (t) is the amount of the substrate j (j=GLU, CN, NH4, O2) in the spatial cell of coordinates x, y and z. After redistribution of the substrates, input of oxygen into the medium is taken into account.

The income of molecular oxygen into the medium is assumed to be proportional to the difference between the oxygen concentration at the equilibrium and the actual concentration, as shown in:

$$\Delta_{O2} = v_{O2}\left(s_{O2}^{Eq} - s_{O2}^{xyz}(t)\right) \tag{C.15}$$

where, Δ_{O2} is the O_2 income in a spatial cell during the current time step (pmol/(TS·SC)); v_{O2} is the rate of income of oxygen into the medium (1/TS); and s_{O2}^{Eq} is the maximum amount of oxygen into a spatial cell under equilibrium conditions. Once the oxygen has been solubilized into the medium, Then, s_{O2}^{xyz} is properly actualized according Eq. C.16.

$$\left\{ s_{O2}^{xyz}(t+1) = s_{O2}^{xyz}(t) + \Delta_{O2} \right\}_{x=1,\dots,iq;y=1,\dots,iq;z=1,\dots,iq} \tag{C.16}$$

Sodium thioglycolate action has been modelled with a simple mechanism. Oxygen reduction by thioglycolic acid follows Eq. C.17:

$$2\,HSCH_2COOH + \frac{1}{2}\,O_2 \rightarrow [SCH_2COOH]_2 + H_2O \tag{C.17}$$

The model assumes that as a result of Eq. C.17 a f_{O2} fraction of the inputted O_2 at the current time step is reduced at expense of the sodium thioglycolate pool. As a consequence, if the sodium thioglycolate action is in process, Eq C.16 can be redrawn as:

$$\left\{ s_{O2}^{xyz}(t+1) = s_{O2}^{xyz}(t) + (1 - f_{O2})\,\Delta_{O2} \right\}_{x=1,\dots,iq;y=1,\dots,iq;z=1,\dots,iq} \tag{C.18}$$

C.8 Implemented model verification

A variety of measures and techniques have been used in order to make sure the model was accurately implemented. As a general measure, the INDISIM-*Saccha* simulator has been implemented using Fortran 90 code in order to take advantage of its new features (mainly, free-form source input, intrinsic array functions, dynamic memory allocation, etc.) which allow for a source code simplification, thus, making the code less error prone. Although newer language specifications exist, Fortran 90 is a qualitative leap from previous versions but it nevertheless allows compatibility with existing code. It is also worth noting that, aiming to avoid otherwise potentially hard to find errors, implicit variables declaration has been avoided by using the Implicit None declaration in all the program units.

Internal model logics and behaviour were systematically investigated by collecting both global and individual data through the simulation length and visually testing them within a spreadsheet program. In particular, global variables include: (i) yeast growth curve; (ii) accumulated death individuals curve; (iii) substrates (glucose, organic N, NH_4, and O_2) uptake curves; (iv) main metabolites production curves (released ethanol and CO_2); (v) substrate effective uptake (defined as uptaken substrate divided by structural mass of the population at the beginning of the time step); (vi) population glucose biomass yield; and population growth rate. Individual data were investigated by way of histogram and included: (i) structural mass distribution; and (ii) genealogical age distribution. It is worth pointing out the dual purpose of this practice which facilitates understanding of the model (or modelled system) and of the consequences of the modelling decisions undertaken.

Because it is crucial to maintain mass balances of the chemical elements considered in this design, a central piece of the verification of INDISIM-*Saccha* is a C, N and O mass balance built internally into the simulator. Although not a guarantee of perfect implementation, it allows detection of otherwise hard to find errors. Specifically, the C, N, and O contents within the system at any particular time step have been followed by means of the existing fraction (%) of the element regarding its initial amount at systems initialization adjusted to 6 decimal points.

Many of the simulator subunits used are identical to the ones used and verified in a previous work (see supplementary material of Portell et al., 2014). Newly developed and formaly verified subunits were the metabolism submodel and the simulation of the sodium thioglycolate action. Both submodels were thoroughly tested by randomly setting input variables of the submodel, according to their allowable values, and subsequently reproducing subroutine output by an independent reimplementation within spreadsheet software. A sufficient number of repetitions (2000 runs) allowed exploration of all subunit parts.

References

Alberghina, L., Smeraldi, C., Ranzi, B.M, Porro, D., 1998. Control by nutrients of growth and cell cycle progression in budding Yeast, analyzed by double-tag flow cytometry. *Journal of Bacteriology* 180 (15), 3864-3872.

Eddy A.A., 1980. Some aspects of amino acid transport in yeast. In *Microorganisms and Nitrogen Sources* (ed. J.W.Payne), pp. 35-62. J. Wiley and Sons Ltd., Chichester.

Ginovart, M., Cañadas J.C., 2008. INDISIM-YEAST: an individual-based simulator on a website for experimenting and investigating diverse dynamics of yeast populations in liquid media. *Journal of Industrial Microbiology and Biotechnology* 35(11), 1359-1366. doi: 10.1007/s10295-008-0436-4.

Ginovart, M., López, D., Gras, A., 2005. Individual-based modelling of microbial activity to study mineralization of C and N and nitrification process in soil. *NonLinear Analysis: Real World Applications* 6(4), 773-795. doi: 10.1016/j.nonrwa.2004.12.005

Ginovart, M., Prats, C., Portell, X., Silbert, M., 2011a. Exploring the lag phase and growth initiation of a yeast culture by means of an Individual-based Model. *Food Microbiology* 28(4), 810-817. doi: 10.1016/j.fm.2010.05.004

Ginovart, M., Prats, C., Portell, X., Silbert, M., 2011b. Analysis of the effect of inoculum characteristics on the first stages of a growing yeast population in beer fermentations by means of an Individual-based Model. *Journal of Industrial Microbiology and Biotechnology* 38(1), 153-165. doi: 10.1007/s10295-010-0840-4.

Gras, A., Ginovart, M., Portell, X., Baveye, P.C., 2010. Individual-based modelling of carbon and nitrogen dynamics in soils: parameterization and sensitivity analysis of abiotic components. *Soil Science* 175, 363-374. doi: 10.1097/SS.0b013e3181eda507

Gras, A., Ginovart, M., Valls, J., Baveye, J.C., 2011. Individual-based modelling of carbon and nitrogen dynamics in soils: parameterization and sensitivity analysis of microbial components. *Ecological Modelling* 222, 1998-2010. doi: 10.1016/j.ecolmodel.2011.03.009.

Grimm V., Berger, U., Bastiansen, F., Eliassen S., et al., 2006. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* 198 (1-2), 115-126. doi: 10.1016/j.ecolmodel.2006.04.023.

Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F., 2010. The ODD protocol: a review and first update. *Ecological Modelling* 221 (23), 2760-2768. doi: 10.1016/j.ecolmodel.2010.08.019.

Hayflick, L., 1968. The limited in vitro lifetime of human diploid cell strains. *Experimental Cell Research* 37 (3), 614-636. doi: 10.1016/0014-4827(65)90211-9

Jenkins, C.L., Kennedy, A.I., Hodgson, J.A., Thurston, P., Smart, K.A., 2003. Impact of serial repitching on lager brewing yeast quality. *Journal of the American Society of Brewing Chemists* 61 (1), 1-9.

Madigan, M.T., 2006. *Brock Biology of microorganisms*. 11th ed. Upper Saddle River: Prentice Hall, 2006. ISBN 0-13-196893-9.

Maskell, D.L., Kennedy, A.I., Hodgson, J.A., Katherine, A.S., 2003. Chronological and replicative lifespan of polyploid *Saccharomyces cerevisiae* (syn. *S. pastorianus*). *FEMS Yeast Research* 3 (2), 201-209. doi: 10.1016/S1567-1356(02)00199-X

Mortimer, R.K., Johnston, J.R., 1959. Life span of individual yeast cells. *Nature* 183, 1751-1752. doi: 10.1038/1831751a0

Panek, A.D., 1991. Storage carbohydrates. In *The Yeasts*, 2nd ed., Vol.4. Yeast Organelles (eds A.H. Rose and J.S. Harrison), pp. 655-678. Academic Press, London.

Powell, C.D., Quain, D.E., Smart, K.A., 2003. Chitin scar breaks in aged *Saccharomyces cerevisiae*. *Microbiology* 149 (11), 3129-3137. doi: 10.1099/mic.0.25940-0.

Porro, D., Vai, M., Vanoni, M., Alberghina, L., Hatzis, C., 2009. Analysis and modeling of growing budding yeast populations at the single cell level. *Cytometry A* 75 (2), 114-120. doi: 10.1002/cyto.a.20689.

Portell, X., Gras, A., Ginovart, M., 2014. INDISIM-*Saccha*, an individual-based model to tackle *Saccharomyces cerevisiae* fermentations. *Ecological Modelling* 279, 12-23. doi:10.1016/j.ecolmodel.2014.02.007.

Prats, C., Ferrer, J., Gras, A., Ginovart, M., 2010. Individual-based modelling and simulation of microbial processes: yeast fermentation and multi-species composting. *Mathematical and Computer Modelling of Dynamical Systems* 16 (6), 489-510.

R Core Team, 2012. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rowen, D.W., Meinke, M., LaPorte, D.C., 1992. GLC3 and GHA1 of *Saccharomyces cerevisiae* are allelic and encode the glycogen branching enzyme. *Molecular and Cellular Biology* 12 (1), 22-29.

Sainz, J., Pizarro, F., Pérez-Correa, J.R., Agosin, E., 2003. Modeling of yeast metabolism and process dynamics in batch fermentation. *Biotechnology and Bioengineering* 81 (7), 818-828. doi: 10.1002/bit.10535.

Steinmeyer, D.E., Shuler, M.L., 1989. Structured model for *Saccharomyces cerevisiae*. *Chemical Engineering Science* 44 (9), 2017-2030. doi: 10.1016/0009-2509(89)85138-3.

Walker, G.M., 1998. *Yeast physiology and biotechnology*. Wiley, Chichester.

Yang, K.M., Lee, N.R., Woo, J.M., Choi W., Zimmermann, M., Blank, L.M., Park, J.B., 2012. Ethanol reduces mitochondrial membrane integrity and thereby impacts carbon metabolism of *Saccharomyces cerevisiae*. *FEMS Yeast Research* 12 (6), 675-84.

Zhang, L., Strouthos, C.G., Wang, Z., Deisboeck, T.S., 2009. Simulating brain tumor heterogeneity with a multiscale agent-based model: linking molecular signatures, phenotypes and expansion rate. *Mathematical and Computer Modeling* 49 (1-2), 307-319. doi: 10.1016/j.mcm.2008.05.011.

Bibliography

References cited in the presented contributions and in the appendices are found at the end of the respective sections.

External references

Alberghina, L., G. Mavelli, G. Drovandi, P. Palumbo, S. Pessina, F. Tripodi, P. Coccetti and M. Vanoni (2012). 'Cell growth and cell cycle in *Saccharomyces cerevisiae*: basic regulatory design and protein–protein interaction network'. *Biotechnology Advances* 30.1, 52–72.

Alberghina, L. and D. Porro (1993). 'Quantative flow cytometry: analysis of protein distributions in budding yeast. A mini-review'. *Yeast* 9.8, 815–823.

Alexander, M. A. and T. W. Jeffries (1990). 'Respiratory efficiency and metabolite partitioning as regulatory phenomena in yeasts'. *Enzyme and microbial technology* 12.1, 2–19.

Allman, R., A. C. Hann, A. P. Phillips, K. L. Martin and D. Lloyd (1990). 'Growth of *Azotobacter vinelandii* with correlation of Coulter cell size, flow cytometric parameters, and ultrastructure'. *Cytometry* 11.7, 822–831.

Araujo, P., A. Gras and M. Ginovart (2014). 'An individual-based model for the study of *Paracoccus denitrificans*, a denitrifying bacterium'. In: *Proceedings book of the V International Conference on Environmental, Industrial and Applied Microbiology. Industrial, medical and environmental applications of microorganisms. Current status and trends*. Ed. by A. Mendez-Vilas, p. 28.

Beckman Coulter, Inc. (2011). *LS 13 320 laser diffraction particle size analyzer instructions for use (Ver.: PN B05577AB)*. Available online at https://www.beckmancoulter.com/wsrportal/techdocs?docname=B05577AB.pdf.

Blake, W. J., M. Kærn, C. R. Cantor and J. J. Collins (2003). 'Noise in eukaryotic gene expression'. *Nature* 422.6932, 633–637.

Bluma, A., T. Höpfner, P. Lindner, C. Rehbock, S. Beutel, D. Riechers, B. Hitzmann and T. Scheper (2010). '*In-situ* imaging sensors for bioprocess monitoring: state of the art'. *Analytical and Bioanalytical Chemistry* 398.6, 2429–2438.

Boekhout, T. and C. P. Kurtzman (1996). 'Principles and methods used in yeast classification, and an overview of currently accepted yeast genera'. In: *Nonconventional Yeasts in Biotechnology*. Springer Berlin Heidelberg, pp. 1–81.

Brehm-Stecher, B. F. and E. A. Johnson (2004). 'Single-cell microbiology: tools, technologies, and applications'. *Microbiology and molecular biology reviews* 68.3, 538–559.

Bryan, A. K., A. Goranov, A. Amon and S. R. Manalis (2010). 'Measurement of mass, density, and volume during the cell cycle of yeast'. *Proceedings of the National Academy of Sciences* 107.3, 999–1004.

Bucci, V., N. Majed, F. L. Hellweger and A. Z. Gu (2012). 'Heterogeneity of intracellular polymer storage states in Enhanced Biological Phosphorus Removal (EBPR)— Observation and Modeling'. *Environmental Science & Technology* 46.6, 3244–3252.

Chieco, P., A. Jonker, B. A. de Boer, J. M. Ruijter and C. J. F. Van Noorden (2013). 'Image Cytometry: protocols for 2D and 3D quantification in microscopic images'. *Progress in Histochemistry and Cytochemistry* 47.4, 211–333.

Dens, E. J., K. Bernaerts, A. R. Standaert, J.-U. Kreft and J. F. Van Impe (2005). 'Cell division theory and individual-based modeling of microbial lag: part II. Modeling lag phenomena induced by temperature shifts'. *International Journal of Food Microbiology* 101.3, 319–332.

Devillers, J., H. Devillers, A. Decourtye and P. Aupinel (2010). 'Internet resources for agent-based modelling'. *SAR and QSAR in Environmental Research* 21.3-4, 337–350.

Dickinson, J. R. and M. Schweizer (2004). *Metabolism and molecular physiology of Saccharomyces cerevisiae*. 2nd ed. CRC Press.

EMD Millipore, Corp. (2013). *Muse R cell analyzer user's guide (0110-7895 Rev G)*. Available online at http://www.millipore.com/userguides/tech1/8tut22.

Eshel, G., G. J. Levy, U. Mingelgrin and M. J. Singer (2004). 'Critical evaluation of the use of laser diffraction for particle-size distribution analysis'. *Soil Science Society of America Journal* 68.3, 736–743.

Feldmann, H. (2012). *Yeast: molecular and cell biology*. John Wiley & Sons.

Ferrer, J. (2010). 'Individual-based modelling of *Plasmodium falciparum* erythrocyte infection in *in vitro* cultures'. PhD Thesis. Barcelona: Universitat Politècnica de Catalunya.

Ferrer, J., C. Prats and D. López (2008). 'Individual-based modelling: an essential tool for microbiology'. *Journal of Biological Physics* 34.1-2, 19–37.

Ferrer, J., J. Vidal, C. Prats, J. Valls, E. Herreros, D. López, A. Giró and D. Gargallo (2007). 'Individual-based model and simulation of *Plasmodium falciparum infected* erythrocyte *in vitro* cultures'. *Journal of Theoretical Biology* 248.3, 448–459.

Ferrer-Miralles, N., J. Domingo-Espín, J. L. Corchero, E. Vázquez and A. Villaverde (2009). 'Microbial factories for recombinant pharmaceuticals'. *Microbial Cell Factories* 8.1, 17.

Foladori, P., A. Quaranta and G. Ziglio (2008). 'Use of silica microspheres having refractive index similar to bacteria for conversion of flow cytometric forward light scatter into biovolume'. *Water Research* 42.14, 3757–3766.

Forsburg, S. L. (2005). 'The yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*: models for cell biology research'. *Gravitational and Space Research* 18.2, 3–9.

Ginovart, M. and J. C. Cañadas (2008). 'INDISIM-YEAST: an individual-based simulator on a website for experimenting and investigating diverse dynamics of yeast populations in liquid media'. *Journal of Industrial Microbiology & Biotechnology* 35.11, 1359–1366.

Ginovart, M. (1997). 'Desenvolupament d'una metodologia fonamentada en la simulació discreta per a l'estudi de sistemes biològics d'interès tecnològic i científic'. PhD Thesis. Barcelona: Universitat Politècnica de Catalunya.

Ginovart, M., D. López and A. Gras (2005). 'Individual-based modelling of microbial activity to study mineralization of C and N and nitrification process in soil'. *Nonlinear Analysis: Real World Applications* 6.4, 773–795.

Ginovart, M., D. López and J. Valls (2002a). 'INDISIM, an individual-based discrete simulation model to study bacterial cultures'. *Journal of Theoretical Biology* 214.2, 305–319.

Ginovart, M., D. López, J. Valls and M. Silbert (2002b). 'Individual based simulations of bacterial growth on agar plates'. *Physica A: Statistical Mechanics and its Applications* 305.3–4, 604–618.

Ginovart, M., D. López, J. Valls and M. Silbert (2002c). 'Simulation modelling of bacterial growth in yoghurt'. *International Journal of Food Microbiology* 73.2-3, 415–425.

Ginovart, M., J. Xifré, D. López and M. Silbert (2007). 'INDISIM-YEAST, an individual-based model to study yeast population in batch cultures'. In: *Communicating current research and educational topics and trends in applied microbiology, Microbiology book series*. 1, pp. 401–409.

Godin, M., F. F. Delgado, S. Son, W. H. Grover, A. K. Bryan, A. Tzur, P. Jorgensen, K. Payer, A. D. Grossman, M. W. Kirschner and S. R. Manalis (2010). 'Using buoyant mass to measure the growth of single cells'. *Nature methods* 7.5, 387–390.

Gras, A. (2004). 'INDISIM-SOM un model discret per a l'estudi de la dinàmica de la matèria orgànica i la nitrificació en sòls'. PhD Thesis. Lleida: Universitat de Lleida.

Gras, A., M. Ginovart, X. Portell and P. C. Baveye (2010). 'Individual-based modeling of carbon and nitrogen dynamics in soils: parameterization and sensitivity analysis of abiotic components'. *Soil science* 175.8, 363–374.

Gras, A., M. Ginovart, J. Valls and P. C. Baveye (2011). 'Individual-based modelling of carbon and nitrogen dynamics in soils: parameterization and sensitivity analysis of microbial components'. *Ecological Modelling* 222.12, 1998–2010.

Graumann, P. L. (2006). 'Different genetic programmes within identical bacteria under identical conditions: the phenomenon of bistability greatly modifies our view on bacterial populations'. *Molecular microbiology* 61.3, 560–563.

Grijspeerdt, K., J.-U. Kreft and W. Messens (2005). 'Individual-based modelling of growth and migration of *Salmonella enteritidis* in hens' eggs'. *International Journal of Food Microbiology* 100.1–3, 323–333.

Grimm, V. (1999). 'Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future?' *Ecological Modelling* 115.2–3, 129–148.

Grimm, V., U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz, G. Huse, A. Huth, J. U. Jepsen, C. Jørgensen, W. M. Mooij, B. Müller, G. Pe'er, C. Piou, S. F. Railsback, A. M. Robbins, M. M. Robbins, E. Rossmanith, N. Rüger, E. Strand, S. Souissi, R. A. Stillman, R. Vabø, U. Visser and D. L. DeAngelis (2006). 'A standard protocol for describing individual-based and agent-based models'. *Ecological Modelling* 198.1–2, 115–126.

Grimm, V., U. Berger, D. L. DeAngelis, J. G. Polhill, J. Giske and S. F. Railsback (2010). 'The ODD protocol: a review and first update'. *Ecological Modelling* 221.23, 2760–2768.

Grimm, V. and S. F. Railsback (2005). *Individual-based modeling and ecology*. Princeton university press.

Hellweger, F. L. and V. Bucci (2009). 'A bunch of tiny individuals—Individual-based modeling for microbes'. *Ecological Modelling* 220.1, 8–22.

Hellweger, F. L., N. D. Fredrick and J. A. Berges (2014). 'Age-correlated stress resistance improves fitness of yeast: support from agent-based simulations'. *BMC systems biology* 8.1, 18.

Höpfner, T., A. Bluma, G. Rudolph, P. Lindner and T. Scheper (2010). 'A review of non-invasive optical-based image analysis systems for continuous bioprocess monitoring'. *Bioprocess and Biosystems Engineering* 33.2, 247–256.

Ishii, S., K. Tago and K. Senoo (2010). 'Single-cell analysis and isolation for microbiology and biotechnology: methods and applications'. *Applied microbiology and biotechnology* 86.5, 1281–1292.

Ishtar Snoek, I. S. and H. Yde Steensma (2007). 'Factors involved in anaerobic growth of *Saccharomyces cerevisiae*'. *Yeast* 24.1, 1–10.

Kell, D. B., H. M. Ryder, A. S. Kaprelyants and H. V. Westerhoff (1991). 'Quantifying heterogeneity: flow cytometry of bacterial cultures'. *Antonie van Leeuwenhoek* 60.3-4, 145–158.

KGaA Merck (2013). *Application note: counting yeast cells for brewing and wine industries can be facilitated by the ScepterTM 2.0 handheld automated cell counter*. Tech. rep. Available online at http://www.millipore.com/techpublications/tech1/an3313en.

Kreft, J.-U., G. Booth and J. W. T. Wimpenny (1998). 'BacSim, a simulator for individual-based modelling of bacterial colony growth'. *Microbiology* 144.12, 3275–3287.

Kreft, J.-U., C. M. Plugge, V. Grimm, C. Prats, J. H. J. Leveau, T. Banitz, S. Baines, J. Clark, A. Ros, I. Klapper, C. J. Topping, A. J. Field, A. Schuler, E. Litchman and F. L. Hellweger (2013). 'Mighty small: observing and modeling individual microbes becomes big science'. *Proceedings of the National Academy of Sciences* 110.45, 18027–18028.

Kurtzman, C. P. and J. W. Fell (1998). *The yeasts: a taxonomic study*. Elsevier.

Lachance, M.-A. (2001). 'Yeasts'. In: *eLS*. John Wiley & Sons.

Lixia, Z., G. Mingfu, G. Dongqi, M. Christensen, H. Xujie, L. Hongmei, F. Yingge and F. Shu (2012). 'Preliminary analysis of yeast communities associated with the spontaneous fermentation of Musalais, a traditional alcoholic beverage of southern Xinjiang, China'. *South African Journal of Enology and Viticulture* 33.1, 95–104.

López, D. (1992). 'Simulació de cultius bacterians. Estudi cinètic i termodinàmic'. PhD Thesis. Barcelona: Universitat de Barcelona.

Mager, W. H. and J. Winderickx (2005). 'Yeast as a model for medical and medicinal research'. *Trends in pharmacological sciences* 26.5, 265–273.

Malakar, P. K. and G. C. Barker (2008). 'Estimating single-cell lag times via a Bayesian scheme'. *Applied and Environmental Microbiology* 74.22, 7098–7099.

Maskell, D. L., A. I. Kennedy, J. A. Hodgson and K. A. Smart (2003). 'Chronological and replicative lifespan of polyploid *Saccharomyces cerevisiae* (syn. *S. pastorianus*)'. *FEMS Yeast Research* 3.2, 201–209.

Musat, N., R. Foster, T. Vagner, B. Adam and M. M. M. Kuypers (2012). 'Detecting metabolic activities in single cells, with emphasis on nanoSIMS'. *FEMS microbiology reviews* 36.2, 486–511.

Pin, C. and J. Baranyi (2008). 'Single-cell and population lag times as a function of cell age'. *Applied and Environmental Microbiology* 74.8, 2534–2536.

Piou, C., U. Berger and V. Grimm (2009). 'Proposing an information criterion for individual-based models developed in a pattern-oriented modelling framework'. *Ecological Modelling* 220.17, 1957–1967.

Popescu, G., Y. Park, N. Lue, C. Best-Popescu, L. Deflores, R. R. Dasari, M. S. Feld and K. Badizadegan (2008). 'Optical imaging of cell mass and growth dynamics'. *American Journal of Physiology-Cell Physiology* 295.2, C538–C544.

Porro, D., B. Gasser, T. Fossati, M. Maurer, P. Branduardi, M. Sauer and D. Mattanovich (2011). 'Production of recombinant proteins and metabolites in yeasts'. *Applied Microbiology and Biotechnology* 89.4, 939–948.

Prats, C., J. Ferrer, A. Giró, D. López and J. Valls (2010). 'Spatial properties in individual-based modelling of microbial systems. Study of the composting process'. In: *Modern multidisciplinary applied microbiology*. Wiley−VCH Verlag GmbH & Co. KGaA, pp. 461–465.

Prats, C. (2008). 'Individual-based modelling of bacterial cultures in the study of the lag phase'. PhD Thesis. Barcelona: Universitat Politècnica de Catalunya.

Railsback, S. F. and V. Grimm (2012). *Agent-based and individual-based modeling: a practical introduction*. Princeton University Press.

Ribéreau-Gayon, J., D. Dubourdieu and B. Donèche (1998). *Traité D'œnologie. Tome 1. Microbiologie du vin. Vinifications*. Paris: Dunod.

Rosenfeld, E., B. Beauvoit, M. Rigoulet and J.-M. Salmon (2002). 'Non-respiratory oxygen consumption pathways in anaerobically-grown *Saccharomyces cerevisiae*: evidence and partial characterization'. *Yeast* 19.15, 1299–1321.

Shuler, M. L., R. Aris and H. M. Tsuchiya (1972). 'Hydrodynamic focusing and electronic cell-sizing techniques'. *Applied microbiology* 24.3, 384–388.

Swinnen, I. A. M., K. Bernaerts, E. J. J. Dens, A. H. Geeraerd and J. F. Van Impe (2004). 'Predictive modelling of the microbial lag phase: a review'. *International Journal of Food Microbiology* 94.2, 137–159.

Tack, I. L. M. M., F. Logist, E. Noriega and J. F. Van Impe (2014). 'An individual-based modeling approach to simulate the effects of cellular nutrient competition on *Escherichia coli* K-12 MG1655 colony behavior and interactions in aerobic structured food systems'. *Food Microbiology*. In press.

Thiele, J. C., W. Kurth and V. Grimm (2014). 'Facilitating parameter estimation and sensitivity analysis of agent-based models: a cookbook using NetLogo and R'. *Journal of Artificial Societies and Social Simulation* 17.3, 11.

Tibayrenc, P., L. Preziosi-Belloy, J.-M. Roger and C. Ghommidh (2010). 'Assessing yeast viability from cell size measurements?' *Journal of biotechnology* 149.1-2, 74–80.

Tuite, M. F. and S. G. Oliver (1991). *Saccharomyces*. Biotechnology Handbooks. New York: Plenum Press.

Turner, J. J., J. C. Ewald and J. M. Skotheim (2012). 'Cell size control in yeast'. *Current Biology* 22.9, R350–R359.

Walker, A. and J. Parkhill (2008). 'Single-cell genomics'. *Nature Reviews Microbiology* 6.3, 176–177.

Walker, G. M. (1998). *Yeast physiology and biotechnology*. John Wiley & Sons.

Wang, Z., L. Millet, M. Mir, H. Ding, S. Unarunotai, J. Rogers, M. U. Gillette and G. Popescu (2011). 'Spatial light interference microscopy (SLIM)'. *Optics express* 19.2, 1016–1026.

Wessel, A. K., L. Hmelo, M. R. Parsek and M. Whiteley (2013). 'Going local: technologies for exploring bacterial microenvironments'. *Nature Reviews Microbiology* 11.5, 337–348.

Wilensky, U. (1999). *NetLogo*. Northwestern University. Evanston, IL: Center for Connected Learning and Computer-Based Modeling.

Personal contributions

Ginovart, M., C. Prats, X. Portell and M. Silbert (2011a). 'Analysis of the effect of inoculum characteristics on the first stages of a growing yeast population in beer fermentations by means of an individual-based model'. *Journal of Industrial Microbiology & Biotechnology* 38.1, 153–165.

Ginovart, M., C. Prats, X. Portell and M. Silbert (2011b). 'Exploring the lag phase and growth initiation of a yeast culture by means of an individual-based model'. *Food Microbiology* 28.4, 810–817.

Portell, X., M. Ginovart, R. Carbó and J. Vives-Rego (2011a). 'Differences in stationary-phase cells of a commercial *Saccharomyces cerevisiae* wine yeast grown in aerobic and

microaerophilic batch cultures assessed by electric particle analysis, light diffraction and flow cytometry'. *Journal of Industrial Microbiology & Biotechnology* 38.1, 141–151.

Portell, X., M. Ginovart, R. Carbo, A. Gras and J. Vives-Rego (2011b). 'Population analysis of a commercial *Saccharomyces cerevisiae* wine yeast in a batch culture by electric particle analysis, light diffraction and flow cytometry'. *FEMS Yeast Research* 11.1, 18–28.

Portell, X., A. Gras, R. Carbó and M. Ginovart (2014a). 'Expanding INDISIM-*Saccha* metabolism to tackle aerobic *Saccharomyces cerevisiae* cultures'. *Ecological Modelling*. Submitted.

Portell, X., A. Gras and M. Ginovart (2014b). 'INDISIM-*Saccha*, an individual-based model to tackle *Saccharomyces cerevisiae* fermentations'. *Ecological Modelling* 279, 12–23.

Portell, X., A. Gras, C. Prats and M. Ginovart (2014c). 'The microbial individual-based model INDISIM-YEAST ready to be used in the free access NetLogo modelling environment'. In: *Industrial, medical and environmental applications of microorganisms: current status and trends*. Ed. by A. Méndez-Vilas. Wageningen Academic Publishers, p. 686.